

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:13:41 ; Search time 269 Seconds
(without alignments)
12.002 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 313949

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	9	2	AAW99195
2	36	94.7	9	2	AAW99196
3	36	94.7	9	2	AAW99197
4	36	94.7	9	2	AAW97572
5	36	94.7	9	2	AAW97375
6	36	94.7	9	2	AAW97374
7	36	94.7	9	8	ADH40333
8	36	94.7	9	8	ADH40334
9	23	60.5	9	5	AAU71211
10	23	60.5	9	5	AAU71428
11	22	57.9	9	3	AAH80162
12	22	57.9	9	6	AAU09449
13	22	57.9	9	6	ABP75298
14	22	57.9	9	6	ABP75138
15	22	57.9	9	6	ABP75181
16	21	55.3	6	3	AAH12035
17	21	55.3	6	5	AAH80810
18	21	55.3	6	7	ADH84629
19	21	55.3	8	3	AAH12089
20	21	55.3	8	7	ADH84662
21	21	55.3	9	2	AAW7427
22	21	55.3	9	2	AAW97373
23	21	55.3	9	3	AAH12092
24	21	55.3	9	3	AAH12086
25	21	55.3	9	4	AAH12506

26	21	55.3	9	4	ABH13927	Abh13927 Human C35
27	21	55.3	9	4	ABH12554	Abh12554 Human C35
28	21	55.3	9	4	ABH13072	Abh13072 Human C35
29	21	55.3	9	4	ABH13115	Abh13115 Human C35
30	21	55.3	9	4	ABH13184	Abh13184 Human C35
31	21	55.3	9	4	ABH13667	Abh13667 Human C35
32	21	55.3	9	4	ABH14131	Abh14131 Human C35
33	21	55.3	9	4	ABH12498	Abh12498 Human C35
34	21	55.3	9	4	ABH13144	Abh13144 Human C35
35	21	55.3	9	4	ABH14390	Abh14390 Human C35
36	21	55.3	9	4	ABH14414	Abh14414 Human C35
37	21	55.3	9	4	ABH12651	Abh12651 Human C35
38	21	55.3	9	4	ABH13415	Abh13415 Human C35
39	21	55.3	9	4	ABH13491	Abh13491 Human C35
40	21	55.3	9	4	ABH12658	Abh12658 Human C35
41	21	55.3	9	4	ABH12774	Abh12774 Human C35
42	21	55.3	9	4	ABH12808	Abh12808 Human C35
43	21	55.3	9	4	ABH12983	Abh12983 Human C35
44	21	55.3	9	4	ABH12995	Abh12995 Human C35
45	21	55.3	9	4	ABH13206	Abh13206 Human C35

ALIGNMENTS

RESULT 1						
AAW99195	ID	AAW99195	strand:	peptide:	9	AA.
XX	XX	AAW99195;				
XX	XX	20-MAY-1999	(first entry)			
DE	XX	Minor histocompatibility antigen HA-1 T-cell epitope #1.				
KW	XX	Minor histocompatibility antigen, HA-1; T-cell epitope; immunological;				
KW	XX	graft versus host disease; bone marrow transplant; leukaemia; vaccine;				
KW	XX	diagnosis; aplastic anaemia; immune deficiency disease.				
OS	XX	Homo sapiens.				
OS	XX	Synthetic.				
FT	XX	Key	Location/Qualifiers			
FT	XX	Misc-difference 3	/label= His, Arg			
FT	XX	W09905174-A1.				
PD	XX	04-FEB-1999.				
PF	XX	23-JUL-1998;	98WO-NL000425.			
PR	XX	23-JUL-1997;	97EP-00202303.			
XX	XX	(UYLE-) RICKSONIV LEIDEN.				
PI	XX	Goulmy EAJM, Hunt DF, Engelhard VH;				
DR	XX	WPI, 1999-153312/13.				
XX	XX	A new minor histocompatibility antigen, HA-1 - useful to treat immune				
XX	XX	diseases and prevent rejection and host versus graft disease in bone				
XX	XX	marrow and organ transplantation.				
PS	XX	Claim 1; Page 32; 47pp; English.				
XX	XX	The present sequence represents a new peptide (P1) constituting a T-cell				
XX	XX	epitope obtainable from the minor histocompatibility antigen HA-1. The				
XX	XX	peptide is immunogenic and can be used as part of a vaccine. P1 is used				
XX	XX	as a medicine, to induce tolerance for transplants, prevent rejection				
XX	XX	and/or graft versus host disease, or to treat (auto) immune diseases. In				
XX	XX	particular it can be used with bone marrow transplantation, in the				
XX	XX	treatment of severe aplastic anaemia, leukaemia, and immune deficiency				

CC	diseases
XX	
SO	Sequence 9 AA;
OY	Query Match
	Best Local Similarity 94.7%; Score 36; DB 2; Length 9;
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy	1 VLXDDLEA 9
Dd	1 VLXDDLEA 9
XX	
RESULT 2	
ID	AAW9196 standard; peptide; 9 AA.
AC	AAW9196
XX	
DT	20-MAY-1999 (first entry)
DE	Minor histocompatibility antigen HA-1 T-cell epitope #2.
KW	Minor histocompatibility antigen, HA-1; T-cell epitope; immunological;
KM	grat versus host disease; bone marrow transplant; leukaemia; vaccine;
OS	diagnosis; aplastic anaemia; immune deficiency disease.
XX	Homo sapiens.
XX	
PN	WO9905174-A1.
PD	04-FEB-1999.
Pf	23-JUL-1998; 98WO-NL000425.
PR	23-JUL-1997; 97EP-00202303.
PA	(UYLE-) RIJXSUNIV LEIDEN.
PI	Goulmy EAJM, Hunt DF, Engelhard VH;
DR	WPI; 1999-153312/13.
PT	A new minor histocompatibility antigen, HA-1 - useful to treat immune
PT	diseases and prevent rejection and host versus grat disease in bone
PT	marrow and organ transplantation.
PS	Claim 3; Page 32; 47pp; English.
CC	The present sequence represents a new peptide (P1) constituting a T-cell
CC	epitope obtainable from the minor histocompatibility antigen HA-1. The
CC	peptide is immunogenic and can be used as part of a vaccine. P1 is used
CC	as a medicine, to induce tolerance for transplants, prevent rejection
CC	and/or graft versus host disease, or to treat (auto) immune diseases. In
CC	particular it can be used with bone marrow transplantation, in the
CC	treatment of severe aplastic anaemia, leukaemia, and immune deficiency
CC	diseases
SQ	Sequence 9 AA;
OY	Query Match
	Best Local Similarity 94.7%; Score 36; DB 2; Length 9;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Dy	1 VLXDDLEA 9
Dd	1 VLHDDLEA 9
XX	
RESULT 3	
ID	AAW9197 standard; peptide; 9 AA.
XX	

AC	AAM99197;
XX	
DT	20-MAY-1999 (first entry)
DE	Minor histocompatibility antigen HA-1 T-cell epitope #3.
XX	
KW	Minor histocompatibility antigen; HA-1; T-cell epitope; immunological; graft versus host disease; bone marrow transplant; leukaemia; vaccine; diagnosis; aplastic anaemia; immune deficiency disease.
XX	
OS	Homo sapiens.
XX	
PN	WO9905174-A1.
XX	
PD	04-FEB-1999.
XX	
PE	23-JUL-1998; 98MO-NL000425.
XX	
PR	23-JUL-1997; 97BP-00202303.
XX	
PA	(UYLE-) RIJXSUNIV LEIDEN.
XX	
PI	Goulmy EAJM, Hunt DF, Engelhard VH;
XX	
DR	WPI; 1999-153312/13.
XX	
PT	A new minor histocompatibility antigen, HA-1 - useful to treat immune diseases and prevent rejection and host versus graft disease in bone marrow and organ transplantation.
XX	
PS	Disclosure; Page 15; 47pp; English.
XX	
CC	The present sequence represents a new peptide (P1) constituting a T-cell epitope obtainable from the minor histocompatibility antigen HA-1. The peptide is immunogenic and can be used as part of a vaccine. P1 is used as a medicing, to induce tolerance for transplants, prevent rejection and/or graft versus host disease, or to treat (auto) immune diseases. In particular it can be used with bone marrow transplantation, in the treatment of severe aplastic anaemia, leukaemia, and immune deficiency diseases
XX	
SQ	Sequence 9 AA;
XX	
Query Match	94.7%; Score 36; DB 2; Length 9;
Best Local Similarity	88.9%; Pred. No. 1.7e+06;
Matches	8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 VLXDDLPEA 9 1 VLRDDLEEA 9
Db	
RESULT 4	
ID	AAM97572 standard; peptide; 9 AA.
XX	
AC	AAM97572;
XX	
DT	20-MAY-1999 (first entry)
XX	
DE	T-cell epitope from the minor histocompatibility antigen HA-1.
XX	
KW	T-cell epitope; minor histocompatibility antigen HA-1; vaccine; transplant rejection; Graft-versus-host Disease; autoimmune disease; neoplastic haematopoietic cell.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 3 /note= "His or Arg"
XX	
PN	WO9905173-A1

```

XX 04-FEB-1999.
PD
XX 23-JUL-1998; 98WO-NL000424.
PF
XX 23-JUL-1997; 97EP-00202303.
PR
XX (UYLE-) RIKXSUNIV LEIDEN.
PA
XX Goulmy EAJM, Hunt DF, Engelhard VH;
PI
XX WPI; 1999-142855/12.
DR
XX Immunogenic peptide from minor histocompatibility antigen HA-1 - useful
PT for inducing tolerance to transplants and prevent rejection or graft-
PT versus-host disease.
XX
XX Claim 1; Page 39, 57pp; English.
PS
XX The present sequence represents an immunogenic peptide constituting a T-
XX cell epitope, obtainable from the minor histocompatibility antigen HA-1.
XX The peptide can be used in vaccines or pharmaceutical formulations as a
XX medicines to induce tolerance for transplants so as to prevent rejection
XX and/or Graft-versus-Host Disease, or to treat autoimmune diseases.
XX Neoplastic haematopoietic cells presenting the peptides, in an HLA class
XX I context, can be eliminated after specific recognition of the peptides.
XX The peptides can also be used to raise antibodies, T-cell receptor, B-
XX and T-cells
XX
XX Sequence 9 AA;
SQ

```

```

Query Match          94.7%; Score 36; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VLXDDLLEA 9
   |||||
DB 1 VLXDDLLEA 9

```

```

RESULT 5
AAW97375
ID AAW97375 standard; protein; 9 AA.
XX
XX AAW97375;
AC
XX
XX 13-MAY-1999 (first entry)
DT
XX
XX HA-1 H-allele sequence.
DE
XX
XX Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
XX R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
XX severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO9905313-A2.
PN
XX
XX 04-FEB-1999.
PD
XX
XX 23-JUL-1998; 98WO-EP004928.
PF
XX
XX 23-JUL-1997; 97EP-00202303.
PR
XX 02-JUN-1998; 98EP-00870125.
XX
XX (UYLE-) RIKXSUNIV LEIDEN.
PA
XX
XX Goulmy E;
PI
XX
XX WPI; 1999-142960/12.
DR
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection

```

```

PT of genetic aberrances.
XX
XX Claim 18; Fig 5; 59pp; English.
PS
XX
XX The present sequence represents part of the minor histocompatibility
CC antigen HA-1 H-allele. The specification describes methods for typing
CC alleles (preferably the H and R alleles) of the minor histocompatibility
CC antigen HA-1 in a sample, which comprise detecting polymorphic
CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC methods can be used for HA-1 typing for bone marrow transplants, severe
CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC detection of genetic aberrances. The probes and primers of the invention
CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
XX anti-idiotypic B cells and/or T cells and antibodies
XX
XX Sequence 9 AA;
SQ

```

```

Query Match          94.7%; Score 36; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 VLXDDLLEA 9
   |||||
DB 1 VLXDDLLEA 9

```

```

RESULT 6
AAW97374
ID AAW97374 standard; protein; 9 AA.
XX
XX AAW97374;
AC
XX
XX 13-MAY-1999 (first entry)
DT
XX
XX HA-1 R-allele sequence.
DE
XX
XX Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
XX R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
XX severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO9905313-A2.
PN
XX
XX 04-FEB-1999.
PD
XX
XX 23-JUL-1998; 98WO-EP004928.
PF
XX
XX 23-JUL-1997; 97EP-00202303.
PR
XX 02-JUN-1998; 98EP-00870125.
XX
XX (UYLE-) RIKXSUNIV LEIDEN.
PA
XX
XX Goulmy E;
PI
XX
XX WPI; 1999-142960/12.
DR
XX
XX Typing minor histocompatibility antigen HA-1 - by amplifying and
PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
PT of genetic aberrances.
XX
XX Claim 13; Fig 5; 59pp; English.
PS
XX
XX The present sequence represents part of the minor histocompatibility
CC antigen HA-1 R-allele. The specification describes methods for typing
CC alleles (preferably the H and R alleles) of the minor histocompatibility
CC antigen HA-1 in a sample, which comprise detecting polymorphic
CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC methods can be used for HA-1 typing for bone marrow transplants, severe
CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC detection of genetic aberrances. The probes and primers of the invention
CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
XX anti-idiotypic B cells and/or T cells and antibodies
XX
XX

```

XX SQ Sequence 9 AA;
Query Match 94.7%; Score 36; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLXDDLLEA 9
DB 1 VLRDDLLLEA 9
RESULT 7
ADH40333 standard; peptide; 9 AA.
XX ADH40333;
AC ADH40333;
XX 11-MAR-2004 (first entry)
DT Human minor histocompatibility antigen HA-1 T cell epitope.
XX Human minor histocompatibility antigen HA-1 T cell epitope.
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
KW minor histocompatibility antigen; mHAg; T cell epitope.
XX Homo sapiens.
OS WO2003106692-A2.
XX PN 24-DEC-2003.
XX PD 13-JUN-2003; 2003WO-EP006251.
PF 13-JUN-2002; 2002EP-00013423.
XX PR 13-JUN-2002; 2002EP-00013423.
XX (MERE) MERCK PATENT GMBH.
XX PA Strltmatter W, Moll H;
XX PI Strltmatter W, Moll H;
XX DR WPI; 2004-082200/08.
XX PT Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX PS Disclosure; Page 82; 119pp; English.
XX CC The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC cancer. This is useful for diagnosing a disease in an individual by
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX SQ Sequence 9 AA;
Query Match 94.7%; Score 36; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLXDDLLEA 9
DB 1 VLRDDLLLEA 9

DB 1 VLHDDLLEA 9
RESULT 8
ADH40334 standard; peptide; 9 AA.
XX ADH40334;
AC ADH40334;
XX 11-MAR-2004 (first entry)
DT Human minor histocompatibility antigen HA-1 T cell epitope.
XX Human minor histocompatibility antigen HA-1 T cell epitope.
XX human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
KW minor histocompatibility antigen; mHAg; T cell epitope.
XX Homo sapiens.
OS WO2003106692-A2.
XX PN 24-DEC-2003.
XX PD 13-JUN-2003; 2003WO-EP006251.
PF 13-JUN-2002; 2002EP-00013423.
XX PR 13-JUN-2002; 2002EP-00013423.
XX (MERE) MERCK PATENT GMBH.
XX PA Strltmatter W, Moll H;
XX PI Strltmatter W, Moll H;
XX DR WPI; 2004-082200/08.
XX PT Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX PS Disclosure; Page 82; 119pp; English.
XX CC The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC cancer. This is useful for diagnosing a disease in an individual by
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX SQ Sequence 9 AA;
Query Match 94.7%; Score 36; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLXDDLLEA 9
DB 1 VLRDDLLLEA 9
RESULT 9
AAU71211 standard; peptide; 9 AA.
XX AAU71211

AC AAU71211;
 XX 26-FEB-2002 (first entry)
 DT
 XX
 DE Human MHC class I molecule HLA-A2 binding 103P3E8 peptide #18.
 XX
 XX 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
 KW tumor; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
 KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
 KW chromosome 9q13-q21.
 XX
 OS Homo sapiens.
 XX
 XX WO200179557-A2.
 PN
 XX
 XX 25-OCT-2001.
 PD
 XX
 XX 12-APR-2001; 2001WO-US012181.
 PF
 XX
 XX 12-APR-2000; 2000US-0196647P.
 PR
 XX
 XX (UROG-) UROGENESYS INC.
 PA
 XX Paris M, Chailita-Bid PM, Raitano AB, Mitchell SC, Afar DEH;
 PI Jakobovits A;
 XX
 DR WPI; 2002-061976/08.
 XX
 XX Monitoring 103P3E8 gene products in sample from patient (suspected of)
 PT having cancer, useful for diagnosing, managing or treating cancers, e.g.
 PT prostate cancer, comprises determining presence of aberrant 103P3E8 gene
 PT products.
 PS
 XX
 XX Disclosure; Page 84; 128pp; English.
 XX
 CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
 CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
 CC expression in normal adult tissue, but it is also aberrantly expressed in
 CC many cancers including tumors of the prostate, bladder, kidney, colon,
 CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
 CC protein and peptide fragments and specific PCR primers are therefore
 CC useful for diagnosing and treating cancer. A vector comprising a
 CC polynucleotide which encodes a single chain monoclonal antibody, that
 CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
 CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
 CC are both useful in the preparation of a composition for treating a
 CC patient with a cancer that expresses 103P3E8. The sequences can be used
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in
 CC serum, blood, urine and tissue and to thereby detect the presence of
 CC cancerous cells
 CC
 XX
 SO Sequence 9 AA;
 SQ

Query Match 60.5%; Score 23; DB 5; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLEA 9
 |||||
 DB 1 DLEA 5

RESULT 10
 AAU71428
 ID AAU71428 standard; peptide; 9 AA.
 XX
 XX AAU71428;
 AC
 XX
 XX 26-FEB-2002 (first entry)
 DT
 XX
 XX Human MHC molecule HLA-A11 binding 103P3E8 peptide #35.
 DE
 XX
 XX 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;

KW tumor; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
 KW single chain monoclonal antibody; serum; blood; urine; tissue; human;
 KW chromosome 9q13-q21.
 XX
 OS Homo sapiens.
 XX
 XX WO200179557-A2.
 PN
 XX
 XX 25-OCT-2001.
 PD
 XX
 XX 12-APR-2001; 2001WO-US012181.
 PF
 XX
 XX 12-APR-2000; 2000US-0196647P.
 PR
 XX
 XX (UROG-) UROGENESYS INC.
 PA
 XX Paris M, Chailita-Bid PM, Raitano AB, Mitchell SC, Afar DEH;
 PI Jakobovits A;
 XX
 DR WPI; 2002-061976/08.
 XX
 XX Monitoring 103P3E8 gene products in sample from patient (suspected of)
 PT having cancer, useful for diagnosing, managing or treating cancers, e.g.
 PT prostate cancer, comprises determining presence of aberrant 103P3E8 gene
 PT products.
 PS
 XX
 XX Disclosure; Page 90; 128pp; English.
 XX
 CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and
 CC peptide fragments of the protein. 103P3E8 exhibits tissue specific
 CC expression in normal adult tissue, but it is also aberrantly expressed in
 CC many cancers including tumors of the prostate, bladder, kidney, colon,
 CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
 CC protein and peptide fragments and specific PCR primers are therefore
 CC useful for diagnosing and treating cancer. A vector comprising a
 CC polynucleotide which encodes a single chain monoclonal antibody, that
 CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme
 CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
 CC are both useful in the preparation of a composition for treating a
 CC patient with a cancer that expresses 103P3E8. The sequences can be used
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in
 CC serum, blood, urine and tissue and to thereby detect the presence of
 CC cancerous cells
 CC
 XX
 SO Sequence 9 AA;
 SQ

Query Match 60.5%; Score 23; DB 5; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLEA 9
 |||||
 DB 5 DLEA 9

RESULT 11
 AA80162
 ID AA80162 standard; peptide; 9 AA.
 XX
 XX AA80162;
 AC
 XX
 XX 24-MAY-2000 (first entry)
 DT
 XX
 XX HLA-A2 restricted tumour antigen peptide derived from SART-1 #15.
 DE
 XX
 XX HLA-A2 restricted tumour antigen; SART-1; human leukocyte antigen; human;
 KW diagnosis; tumour; cytotoxic T cell; flat epithelioma; lung cancer;
 KW oesophagus cancer; cytostatic.
 KW
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200006595-A1.
 PN

XX 10-FEB-2000.
 XX 27-JUL-1999; 99WO-JP004010.
 XX 28-JUL-1998; 98JP-00212940.
 XX (SUMU) SUMITOMO PHARM CO LTD.
 XX (ITOH/) ITOH K.
 XX Itoh K, Kobayashi T;
 XX WPI; 2000-195258/17.
 XX HLA-A2 resistant tumor antigen peptides derived from SART-1, useful as
 XX preventives or diagnostics for tumors e.g. flat epithelioma like lung
 XX cancer.
 XX Claim 2; Page 41; 55pp; Japanese.
 CC AAY80148 to AAY80181 represent specifically claimed HLA-A2 restricted
 CC tumour antigen peptides derived from SART-1. The peptides have cytostatic
 CC activity. The peptides are useful as a preventive or diagnostic for
 CC tumours e.g. flat epithelioma like lung cancer and oesophagus cancer
 CC Sequence 9 AA;
 SQ

Query Match 57.9%; Score 22; DB 3; Length 9;
 Best Local Similarity 71.4%; Pred. No. 1.7e+06; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 2;

OY 1 VLXDDL 7
 | | | | |
 DB 3 VRADDL 9

RESULT 12
 AAU09449 ID AAU09449 standard; peptide; 9 AA.
 AC AAU09449;

DT 26-MAR-2002 (first entry)

DE Chlamydia pneumoniae T-cell epitope #4 useful as Chlamydia antigen.
 XX
 KM ATP binding cassette; secretory locus open reading frame; endopeptidase;
 KM secretory locus ORF; protease; metalloprotease; CLP protease ATPase;
 KM CLP protease subunit; transglycolase/transpeptidase; CLPC protease;
 KM thioredoxin; Chlamydia infection; antibacterial; immunogen.
 XX

OS Chlamydia pneumoniae CML029.

XX WO200185972-A2.

PN 15-NOV-2001.

PF 08-MAY-2001; 2001WO-CA000653.

PR 08-MAY-2000; 2000US-0202672P.

PR 30-MAY-2000; 2000US-0207852P.

PR 16-JUN-2000; 2000US-0211796P.

PR 16-JUN-2000; 2000US-0211797P.

PR 16-JUN-2000; 2000US-0211798P.

PR 16-JUN-2000; 2000US-0211801P.

PR 16-JUN-2000; 2000US-0212044P.

PR 26-SEP-2000; 2000US-0235335P.

PR 26-SEP-2000; 2000US-0235361P.

PR 26-SEP-2000; 2000US-0235398P.

(AVERT) AVENTIS PASTEUR LTD.

PI Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2002-049447/06.
 XX Vaccine useful for immunizing mammals against chlamydia infections,
 XX comprises vectors having sequences of ATP binding cassette gene,
 XX secretory locus open reading frame gene of chlamydia.
 XX Example 4; Page 80; 355pp; English.

XX The present invention relates to the isolation of Chlamydia pneumoniae
 XX pneumoniae strain CML029 genes and their encoded proteins. The genes of
 XX the invention encode an ATP binding cassette gene, a secretory locus open
 XX reading frame (ORF), an endopeptidase, a protease, a metalloprotease, CLP
 XX protease ATPase, or thioredoxin. The genes of the invention can be used
 XX in a vector as a vaccine for the prevention and treatment of Chlamydia
 XX infections. AAU09440-AAU09473 represent B- or T-cell epitopes from the C.
 XX pneumoniae proteins (AAU09430-AAU09439) of the invention. These epitopes
 XX can be used as Chlamydia antigens
 XX

SQ Sequence 9 AA;

Query Match 57.9%; Score 22; DB 5; Length 9;
 Best Local Similarity 57.1%; Pred. No. 1.7e+06; Indels 0; Gaps 0;
 Matches 4; Conservative 2; Mismatches 1;

OY 2 LXXDLL 8
 | | | | |
 DB 2 LGDRLE 8

RESULT 13
 ABP75298 ID ABP75298 standard; peptide; 9 AA.
 AC ABP75298;

DT 20-FEB-2003 (first entry)

DE Chlamydia pneumoniae peptide epitope #64.

XX Antibacterial; secreted protein; intracellular bacterium.

OS Chlamydia pneumoniae.

PN WO200282091-A2.

PD 17-OCT-2002.

PF 09-APR-2002; 2002WO-DK000234.

PR 09-APR-2001; 2001DK-00000581.

PR 09-APR-2001; 2001US-0282513P.

PA (SHAW/) SHAW A C.

PA (VAND/) VANDAH B B.

PI Shaw AC, Vandahl BB;

DR WPI; 2003-058585/05.

XX Identifying intracellular bacterial proteins by labeling proteins in the
 XX presence of a eukaryotic protein synthesis inhibitor, performing
 XX electrophoresis, autoradiography and comparing profiles to an infected-
 XX cell lysate profile.

PS Claim 34; Page 151; 179pp; English.

XX The present invention relates to a method (M1) for identifying secreted
 XX intracellular bacterial proteins (BP). M1 comprises: (a) selectively
 XX visualising BP by pulse labelling in the presence of an inhibitor of
 XX eukaryotic protein synthesis followed by 2D electrophoresis and
 XX autoradiography; (b) comparing protein profiles (PP) of purified bacteria

CC to PF of total lysate (TL) of infected cells; and (c) identifying protein
CC spots present in differential images from gels loaded with TL. The
CC present sequence is one such bacterial peptide epitope which was
CC identified by the method of the invention

XX Sequence 9 AA;

Query Match 57.9%; Score 22; DB 6; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.7e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXXDLE 8
| | | | |
Db 2 LGDEILE 8

RESULT 14
ABP75138
ID ABP75138 standard; peptide; 9 AA.

XX ABP75138;

DT 03-FEB-2003 (first entry)

DE Proteome analysis related peptide #423.

XX Proteome analysis; isolation; determination; diagnostic assay; detection;
KM protein marker; identification; metastatic; invasive cancer;
XX differential expression; signalling pathway; chromatography.

OS Synthetic.

PN WO200277016-A2.

PD 03-OCT-2002.

PF 22-MAR-2002; 2002WO-EP003368.

PR 22-MAR-2001; 2001US-0278171P.

PR 12-SEP-2001; 2001US-0318749P.

PR 20-SEP-2001; 2001US-0323999P.

PA (VLA-A-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Vandekerckhove J, Gevaert K;

DR WPI; 2003-067379/06.

PT Method for isolation of peptides from complex mixture of peptides
PT involves specific chemical and/or enzymatic alteration of at least one
PT type of peptide.

PS Example 22; Page 150; 193pp; English.

XX The present invention describes a method (M1) for the isolation of a
CC subset of peptides from a protein peptide mixture (P1). M1 involves: (a)
CC separating the protein peptide mixture into fractions of peptides via
CC chromatography; (b) chemically, or enzymatically, or chemically and
CC enzymatically, altering at least one amino acid of at least one of the
CC peptides in each fraction, thereby generating a subset of altered
CC peptides; and (c) isolating the altered (flagged) peptides out of each
CC fraction via chromatography, where the chromatography of steps (a) and
CC (c) is performed with the same type of chromatography. M1 can be used for
CC the isolation and determination of peptides from protein peptide
CC mixtures. M1 can also be used in diagnostic assays for detection of the
CC presence, the absence or a variation in expression level of at least one
CC protein marker or a specific set of proteins indicative of a disease
CC state. M1 can be used for identifying target proteins present in
CC metastatic and invasive cancers, in differential expression of proteins
CC in transgenic mice, identification of proteins that are upregulated or
CC down regulated in disease tissues, in identification of intracellular
CC changes in cells with physiological changes such as metabolic shift, in
CC the identification of biomarkers in cancers and in the identification of

CC signalling pathways. The method is gel-free methodology for qualitative
CC and quantitative proteome analysis without the need for multidimensional
CC chromatography and without the use of affinity tags. ABP74714 to ABP75190
CC represent peptide sequences used in the exemplification of the present
CC invention

XX Sequence 9 AA;

Query Match 57.9%; Score 22; DB 6; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 LXXDLE 8
| | | | |
Db 3 LLDLSE 9

RESULT 15
ABP75181
ID ABP75181 standard; peptide; 9 AA.

XX ABP75181;

DT 03-FEB-2003 (first entry)

DE Proteome analysis related peptide #466.

XX Proteome analysis; isolation; determination; diagnostic assay; detection;
KM protein marker; identification; metastatic; invasive cancer;
XX differential expression; signalling pathway; chromatography.

OS Synthetic.

PN WO200277016-A2.

PD 03-OCT-2002.

PF 22-MAR-2002; 2002WO-EP003368.

PR 22-MAR-2001; 2001US-0278171P.

PR 12-SEP-2001; 2001US-0318749P.

PR 20-SEP-2001; 2001US-0323999P.

PA (VLA-A-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PI Vandekerckhove J, Gevaert K;

DR WPI; 2003-067379/06.

PT Method for isolation of peptides from complex mixture of peptides
PT involves specific chemical and/or enzymatic alteration of at least one
PT type of peptide.

PS Example 22; Page 151; 193pp; English.

XX The present invention describes a method (M1) for the isolation of a
CC subset of peptides from a protein peptide mixture (P1). M1 involves: (a)
CC separating the protein peptide mixture into fractions of peptides via
CC chromatography; (b) chemically, or enzymatically, or chemically and
CC enzymatically, altering at least one amino acid of at least one of the
CC peptides in each fraction, thereby generating a subset of altered
CC peptides; and (c) isolating the altered (flagged) peptides out of each
CC fraction via chromatography, where the chromatography of steps (a) and
CC (c) is performed with the same type of chromatography. M1 can be used for
CC the isolation and determination of peptides from protein peptide
CC mixtures. M1 can also be used in diagnostic assays for detection of the
CC presence, the absence or a variation in expression level of at least one
CC protein marker or a specific set of proteins indicative of a disease
CC state. M1 can be used for identifying target proteins present in
CC metastatic and invasive cancers, in differential expression of proteins
CC in transgenic mice, identification of proteins that are upregulated or
CC down regulated in disease tissues, in identification of intracellular
CC changes in cells with physiological changes such as metabolic shift, in

CC the identification of biomarkers in cancers and in the identification of
 CC signalling pathways. The method is gel-free methodology for qualitative
 CC and quantitative proteome analysis without the need for multidimensional
 CC chromatography and without the use of affinity tags. ABP74714 to ABP75190
 CC represent peptide sequences used in the exemplification of the present
 CC invention

XX
 SQ Sequence 9 AA;

Query Match 57.9%; Score 22; DB 6; Length 9;
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKDDLL 8
 | | | | |
 Db 3 LKDDLSB 9

Search completed: December 29, 2004, 23:29:52
 Job time : 274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:24:55 ; Search time 37 Seconds
(without alignments)
16.131 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLERA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 95011

Minimum DB seq length: 0
Maximum DB seq length: 957

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	57.9	7	US-08-556-419-14	Sequence 14, Appl
2	22	57.9	9	US-09-744-549-15	Sequence 15, Appl
3	21	55.3	6	US-08-459-568-20	Sequence 20, Appl
4	21	55.3	6	US-08-399-411-20	Sequence 20, Appl
5	21	55.3	6	US-08-516-859A-20	Sequence 20, Appl
6	21	55.3	6	US-09-586-472-20	Sequence 20, Appl
7	21	55.3	6	US-09-528-706-20	Sequence 20, Appl
8	21	55.3	6	US-08-459-568-75	Sequence 20, Appl
9	21	55.3	8	US-08-399-411-75	Sequence 20, Appl
10	21	55.3	8	US-08-516-859A-75	Sequence 20, Appl
11	21	55.3	8	US-09-528-706-75	Sequence 20, Appl
12	21	55.3	8	US-08-459-568-72	Sequence 20, Appl
13	21	55.3	9	US-08-459-568-78	Sequence 20, Appl
14	21	55.3	9	US-08-399-411-72	Sequence 20, Appl
15	21	55.3	9	US-08-516-859A-72	Sequence 20, Appl
16	21	55.3	9	US-09-528-706-72	Sequence 20, Appl
17	21	55.3	9	US-08-516-859A-72	Sequence 20, Appl
18	21	55.3	9	US-08-516-859A-78	Sequence 20, Appl
19	21	55.3	9	US-09-528-706-78	Sequence 20, Appl
20	21	55.3	9	US-08-873-235B-11	Sequence 11, Appl
21	21	55.3	9	US-09-586-472-72	Sequence 72, Appl
22	21	55.3	9	US-09-586-472-78	Sequence 72, Appl
23	21	55.3	9	US-09-528-706-72	Sequence 72, Appl
24	21	55.3	9	US-09-528-706-78	Sequence 72, Appl
25	20	52.6	6	US-09-023-819-11	Sequence 11, Appl
26	20	52.6	6	US-09-808-126-11	Sequence 11, Appl
27	20	52.6	6	US-09-803-951-11	Sequence 11, Appl

28	20	52.6	8	US-08-747-539A-47	Sequence 47, Appl
29	20	52.6	8	US-08-747-539A-50	Sequence 50, Appl
30	20	52.6	8	US-08-747-539A-51	Sequence 51, Appl
31	19	50.0	6	US-08-974-549A-373	Sequence 373, Appl
32	19	50.0	6	US-08-912-951-140	Sequence 140, Appl
33	19	50.0	6	US-09-402-161B-373	Sequence 373, Appl
34	19	50.0	7	US-09-721-456-373	Sequence 373, Appl
35	19	50.0	7	US-09-187-859-1186	Sequence 1186, Appl
36	19	50.0	7	US-09-839-542B-1186	Sequence 1186, Appl
37	19	50.0	8	US-09-187-859-1189	Sequence 1189, Appl
38	19	50.0	8	US-09-187-859-2550	Sequence 2550, Appl
39	19	50.0	8	US-09-183-266A-35	Sequence 35, Appl
40	19	50.0	8	US-09-183-266A-36	Sequence 36, Appl
41	19	50.0	8	US-09-183-266A-38	Sequence 38, Appl
42	19	50.0	8	US-09-839-542B-1189	Sequence 1189, Appl
43	19	50.0	8	US-09-839-542B-2550	Sequence 2550, Appl
44	19	50.0	9	US-07-663-413-15	Sequence 15, Appl
45	19	50.0	9	US-08-055-530-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-556-419-14
; Sequence 14, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lananhan, Anthony
; APPLICANT: Morley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.52271
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-556-419-14

Query Match 57.9% Score 22; DB 3; Length 7;
Best Local Similarity 80.0% Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8
|||:
Db 1 DDLQ 5

RESULT 2
US-09-744-549-15
; Sequence 15, Application US/09744549
; Patent No. 6664232
; GENERAL INFORMATION:
; APPLICANT: Itoh, Kyogo et al.
; TITLE OF INVENTION: HLA-A2-restricted Tumor Antigen Peptides Derived From SART-1
; FILE REFERENCE: 0020-4808P
; CURRENT APPLICATION NUMBER: US/09/744,549
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: JP H10-212940
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 9
; TYPE: PRT

ORGANISM: Homo sapiens
US-09-744-549-15

Query Match 57.9%; Score 21; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0;

QY 1 VLXDDL 7
| | | | |
Db 3 VRADDL 9

RESULT 3

US-08-459-568-20
Sequence 20, Application US/08459568
Patent No. 581304

GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,568
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-459-568-20
Query Match 55.3%; Score 21; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 4 DDLE 8
: | | | |
Db 1 EDLLE 5

US-08-399-411-20
Sequence 20, Application US/08399411
Patent No. 5831008
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

US-08-399-411-20
Sequence 20, Application US/08399411
Patent No. 5831008
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

US-08-399-411-20
Sequence 20, Application US/08399411
Patent No. 5831008
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,411
FILING DATE: 06-MAR-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-399-411-20
Query Match 55.3%; Score 21; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 4 DDLE 8
: | | | |
Db 1 EDLLE 5

US-08-516-859A-20
Sequence 20, Application US/08516859A
Patent No. 6069231

GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

US-08-516-859A-20
Sequence 20, Application US/08516859A
Patent No. 6069231
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

US-08-516-859A-20
Sequence 20, Application US/08516859A
Patent No. 6069231
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

US-08-516-859A-20
Sequence 20, Application US/08516859A
Patent No. 6069231
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-516-859A-20

Query Match 55.3%; Score 21; DB 3; Length 6;
Beet Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
DB 1 EDLE 5

RESULT 6
US-09-586-472-20
Sequence 20, Application US/09586472
Patent No. 632335
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/586,472
FILING DATE: 01-Jun-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/528,706
FILING DATE: 17-MAR-2000
APPLICATION NUMBER: US 08/516,859
FILING DATE: 18-AUG-1995
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
APPLICATION NUMBER: US 08/292,683
FILING DATE: 18-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 4130

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-586-472-20

Query Match 55.3%; Score 21; DB 3; Length 6;
Beet Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8

DB 1 EDLE 5

RESULT 7
US-09-528-706-20
Sequence 20, Application US/09528706
Patent No. 6468985
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/528,706
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/516,859
FILING DATE:
APPLICATION NUMBER: US 08/292,683
FILING DATE: 18-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-528-706-20

Query Match 55.3%; Score 21; DB 4; Length 6;
Beet Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
DB 1 EDLE 5

RESULT 8
US-08-459-568-75
Sequence 75, Application US/08459568
Patent No. 581304
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,568
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-459-568-75

Query Match 55.3%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLE 8
:||||
Db 1 EDLLE 5

RESULT 9
US-08-399-411-75
Sequence 75, Application US/08399411
Patent No. 5831008
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,411
FILING DATE: 06-MAR-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-399-411-75

Query Match 55.3%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLE 8
:||||
Db 1 EDLLE 5

RESULT 10
US-08-516-859A-75
Sequence 75, Application US/08516859A
Patent No. 6069231
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,683
FILING DATE: 18-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-516-859A-75

Query Match 55.3%; Score 21; DB 3; Length 8;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLE 8
:||||
Db 1 EDLLE 5

RESULT 11
US-09-586-472-75
Sequence 75, Application US/09586472
Patent No. 6323335
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
Zinc Finger Proteins

```

1      NUMBER OF SEQUENCES: 106
2      CORRESPONDENCE ADDRESS:
3      ADDRESSEE: Campbell & Flores LLP
4      STREET: 4370 La Jolla Village Drive, Suite 700
5      CITY: San Diego
6      STATE: California
7      COUNTRY: USA
8      ZIP: 92122
9
10     COMPUTER READABLE FORM:
11     MEDIUM TYPE: Floppy disk
12     COMPUTER: IBM PC compatible
13     OPERATING SYSTEM: PC-DOS/MS-DOS
14     SOFTWARE: PatentIn Release #1.0, Version #1.25
15
16     CURRENT APPLICATION DATA:
17     APPLICATION NUMBER: US/09/586,472
18     FILING DATE: 01-Jun-2000
19     CLASSIFICATION: <Unknown>
20
21     PRIOR APPLICATION DATA:
22     APPLICATION NUMBER: US 09/528,706
23     FILING DATE: 17-MAR-2000
24     APPLICATION NUMBER: US 08/516,859
25     FILING DATE: 18-AUG-1995
26     APPLICATION NUMBER: US 08/399,411
27     FILING DATE: 06-MAR-1995
28     APPLICATION NUMBER: US 08/292,663
29     FILING DATE: 18-AUG-1994
30
31     ATTORNEY/AGENT INFORMATION:
32     NAME: Campbell, Cathryn A.
33     REGISTRATION NUMBER: 31,815
34     REFERENCE/DOCKET NUMBER: P-LJ 4130
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: (619) 535-9001
37     TELEFAX: (619) 535-8949
38
39     INFORMATION FOR SEQ ID NO: 75:
40     SEQUENCE CHARACTERISTICS:
41     LENGTH: 8 amino acids
42     TYPE: amino acid
43     TOPOLOGY: linear
44
45     US-09-586-472-75
46     SEQUENCE DESCRIPTION: SEQ ID NO: 75:
47
48     Query Match      55.3%; Score 21; DB 3; Length 8;
49     Best Local Similarity 80.0%; Pred. NO. 3.8e+05;
50     Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0
51
52     Cq      4 DDLLE 8
53           :||||
54     Db      1 EDLLE 5
55
56     RESULT 12
57     US-09-528-706-75
58     Sequence 75, Application US/09528706
59     Patent No. 6468985
60     GENERAL INFORMATION:
61     APPLICANT: Huang, Shi
62     TITLE OF INVENTION: Retinoblastoma Protein - Interacting
63     TITLE OF INVENTION: Zinc Finger Proteins
64     NUMBER OF SEQUENCES: 106
65     CORRESPONDENCE ADDRESS:
66     ADDRESSEE: Campbell & Flores LLP
67     STREET: 4370 La Jolla Village Drive, Suite 700
68     CITY: San Diego
69     STATE: California
70     COUNTRY: USA
71     ZIP: 92122
72
73     COMPUTER READABLE FORM:
74     MEDIUM TYPE: Floppy disk
75     COMPUTER: IBM PC compatible
76     OPERATING SYSTEM: PC-DOS/MS-DOS
77     SOFTWARE: PatentIn Release #1.0, Version #1.25
78     CURRENT APPLICATION DATA:
79     APPLICATION NUMBER: US/09/528,706

```

1 FILING DATE: 1
2 CLASSIFICATION: 1
3 PRIOR APPLICATION DATA: 1
4 APPLICATION NUMBER: US 08/516,859 1
5 FILING DATE: 1
6 PRIOR APPLICATION DATA: 1
7 APPLICATION NUMBER: US 08/292,683 1
8 FILING DATE: 18-AUG-1994 1
9 ATTORNEY/AGENT INFORMATION: 1
10 NAME: Campbell, Cathryn A. 1
11 REGISTRATION NUMBER: 31,815 1
12 REFERENCE/POCKET NUMBER: P-LJ 1776 1
13 TELECOMMUNICATION INFORMATION: 1
14 TELEPHONE: (619) 535-9001 1
15 TELEFAX: (619) 535-8949 1
16 INFORMATION FOR SEQ ID NO: 75: 1
17 SEQUENCE CHARACTERISTICS: 1
18 LENGTH: 8 amino acids 1
19 TYPE: amino acid 1
20 TOPOLOGY: linear 1
21 US-09-528-706-75 1

Query Match	55.3%;	Score 21;	DB 4;	Length 8;
Best Local Similarity	80.0%;	Pred. No. 3.8e+05;		
Matches	4;	Conservative	1;	Mismatches 0;
			Indels	0;
Gaps				0;
QY	4	DDLE	8	
	:			
db	1	EDLE	5	

QY	4	DDLE	8
	:		
Db	1	EDLE	5

RESULT 13
 US-08-459-568-72
 Sequence 72: Application US/08459568
 Patent No. 581104
 GENERAL INFORMATION:
 APPLICANT: Huang, Shi
 TITLE OF INVENTION: Retinoblastoma Protein - Interacting
 TITLE OF INVENTION: Zinc Finger Proteins
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,568
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/399,411
 FILING DATE: 06-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1264
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 72:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-459-568-72

Query Match 55.3%; Score 21; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLLE 8
:||||
Db 1 EDLLE 5

RESULT 14

US-08-459-568-78
; Sequence 78, Application US/08459568
; Patent No. 5811304
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,568
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-459-568-78

Query Match 55.3%; Score 21; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLLE 8
:||||
Db 1 EDLLE 5

RESULT 15
US-08-399-411-72
; Sequence 72, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego

STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,411
FILING DATE: 06-MAR-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-399-411-72

Query Match 55.3%; Score 21; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLLE 8
:||||
Db 1 EDLLE 5

Search completed: December 29, 2004, 23:34:37
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:33:19 ; Search time 159 Seconds
(without alignments)
20.362 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 35972711 residues
Total number of hits satisfying chosen parameters: 139774

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	36	94.7	9	US-10-623-176-1	Sequence 1, Appl1
2	36	94.7	9	US-10-623-176-2	Sequence 2, Appl1
3	36	94.7	9	US-10-623-176-10	Sequence 10, Appl1
4	36	94.7	9	US-10-791-217-1	Sequence 1, Appl1
5	36	94.7	9	US-10-791-217-2	Sequence 2, Appl1
6	36	94.7	9	US-10-791-217-5	Sequence 5, Appl1
7	32	84.2	9	US-10-623-176-41	Sequence 41, Appl1
8	32	84.2	9	US-10-623-176-42	Sequence 42, Appl1
9	32	84.2	9	US-10-623-176-45	Sequence 45, Appl1
10	32	84.2	9	US-10-623-176-46	Sequence 46, Appl1
11	29	76.3	9	US-10-623-176-47	Sequence 47, Appl1
12	29	76.3	9	US-10-623-176-48	Sequence 48, Appl1
13	27	71.1	9	US-10-623-176-23	Sequence 23, Appl1

14	27	71.1	9	US-10-623-176-40	Sequence 40, Appl1
15	23	60.5	8	US-10-623-176-15	Sequence 15, Appl1
16	23	60.5	9	US-09-834-765-132	Sequence 132, App
17	23	60.5	9	US-09-834-765-349	Sequence 349, App
18	23	60.5	9	US-10-623-176-3	Sequence 3, Appl1
19	23	60.5	9	US-10-623-176-4	Sequence 4, Appl1
20	23	60.5	9	US-10-623-176-5	Sequence 5, Appl1
21	22	57.9	9	US-10-119-536A-109	Sequence 109, App
22	22	57.9	9	US-10-394-980-421	Sequence 421, App
23	22	57.9	9	US-10-394-980-464	Sequence 464, App
24	22	57.9	9	US-10-275-652-50	Sequence 50, Appl1
25	21	55.3	9	US-09-833-203-54	Sequence 54, Appl1
26	21	55.3	9	US-10-057-475B-10608	Sequence 10608, A
27	21	55.3	9	US-10-057-475B-10813	Sequence 10813, A
28	21	55.3	9	US-10-154-884B-10608	Sequence 10608, A
29	21	55.3	9	US-10-154-884B-10813	Sequence 10813, A
30	21	55.3	9	US-10-623-176-14	Sequence 14, Appl1
31	21	55.3	9	US-10-791-217-4	Sequence 4, Appl1
32	21	55.3	9	US-10-809-790-23	Sequence 23, Appl1
33	21	55.3	9	US-10-166-225A-168	Sequence 168, App
34	20	52.6	5	US-10-166-225A-169	Sequence 169, App
35	20	52.6	5	US-10-166-225A-170	Sequence 170, App
36	20	52.6	5	US-10-166-225A-171	Sequence 171, App
37	20	52.6	5	US-10-166-225A-172	Sequence 172, App
38	20	52.6	5	US-10-166-225A-173	Sequence 173, App
39	20	52.6	6	US-09-803-951-11	Sequence 11, Appl1
40	20	52.6	6	US-09-803-951-11	Sequence 11, Appl1
41	20	52.6	7	US-09-943-944E-85	Sequence 85, Appl1
42	20	52.6	8	US-09-801-784-19	Sequence 19, Appl1
43	20	52.6	8	US-09-801-784-20	Sequence 20, Appl1
44	20	52.6	8	US-09-801-784-21	Sequence 21, Appl1
45	20	52.6	8	US-09-801-784-22	Sequence 22, Appl1

ALIGNMENTS

RESULT 1
US-10-623-176-1
Sequence 1, Application US/10623176
Publication No. US2004009246A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhardt, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT FILING DATE: 2003-07-18
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
OTHER INFORMATION: wherein X can be R or H
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-1
Query Match 94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. NO. 1.5e+06;

```
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 2
US-10-623-176-2
; Sequence 2, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-2

Query Match          94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 3
US-10-623-176-10
; Sequence 10, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Els A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 9
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-10

Query Match          94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 4
US-10-791-217-1
; Sequence 1, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Elsa A.J.M
; APPLICANT: Hunt, Donald F
; APPLICANT: Engelhard, Victor H
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from histocompatibility antigen
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa 16 HISTIDINE OR ARGININE RESIDUE
US-10-791-217-1

Query Match          94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 5
US-10-791-217-2
; Sequence 2, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Elsa A.J.M
; APPLICANT: Hunt, Donald F
; APPLICANT: Engelhard, Victor H
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
```


NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: obtained from histocompatibility antigen
US-10-791-217-2

Query Match 94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
|||
Db 1 VLHDDLLEA 9

RESULT 6
US-10-791-217-5
Sequence 5, Application US/10791217
Publication No. US20040191268A1
GENERAL INFORMATION:
APPLICANT: Goulimy, Elsa A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: The HA-1 Antigen
FILE REFERENCE: 2183-4285US
CURRENT APPLICATION NUMBER: US/10/791,217
CURRENT FILING DATE: 2004-03-02
PRIOR APPLICATION NUMBER: US/09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: obtained from KIAA0223 partial complementary DNA
US-10-791-217-5

Query Match 94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
|||
Db 1 VLHDDLLEA 9

RESULT 7
US-10-623-176-41
Sequence 41, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulimy, Elsa A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165

PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-41

Query Match 84.2%; Score 32; DB 15; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLE 8
|||
Db 2 VLHDDLLE 9

RESULT 8
US-10-623-176-42
Sequence 42, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulimy, Elsa A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 42
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-42

Query Match 84.2%; Score 32; DB 15; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLE 8
|||
Db 2 VLHDDLLE 9

RESULT 9
US-10-623-176-45
Sequence 45, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulimy, Elsa A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.

```
/ TITLE OF INVENTION: HA-1 epitopes and uses thereof
/ FILE REFERENCE: 2183-6047US
/ CURRENT APPLICATION NUMBER: US/10/623,176
/ CURRENT FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: 09/489,760
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: EP 97202303.0
/ PRIOR FILING DATE: 1997-07-23
/ PRIOR APPLICATION NUMBER: PCT/NL98/00424
/ PRIOR FILING DATE: 1998-07-23
/ PRIOR APPLICATION NUMBER: JP 2000-504165
/ PRIOR FILING DATE: 2000-01-24
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 45
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)..(9)
US-10-623-176-45
```

```
Query Match      84.2%; Score 32; DB 15; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LXDILLEA 9
        | | | | |
Db      1 LRDDILLEA 8
```

```
RESULT 10
US-10-623-176-46
/ Sequence 46, Application US/10623176
/ Publication No. US20040092446A1
/ GENERAL INFORMATION:
/ APPLICANT: Goulmy, Els A.J.M.
/ APPLICANT: Hunt, Donald F.
/ APPLICANT: Engelhard, Victor H.
/ TITLE OF INVENTION: HA-1 epitopes and uses thereof
/ FILE REFERENCE: 2183-6047US
/ CURRENT APPLICATION NUMBER: US/10/623,176
/ CURRENT FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: 09/489,760
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: EP 97202303.0
/ PRIOR FILING DATE: 1997-07-23
/ PRIOR APPLICATION NUMBER: PCT/NL98/00424
/ PRIOR FILING DATE: 1998-07-23
/ PRIOR APPLICATION NUMBER: JP 2000-504165
/ PRIOR FILING DATE: 2000-01-24
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)..(9)
US-10-623-176-46
```

```
Query Match      84.2%; Score 32; DB 15; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 LXDILLEA 9
        | | | | |
```

```
Db      1 LRDDILLEA 8
```

```
RESULT 11
US-10-623-176-47
/ Sequence 47, Application US/10623176
/ Publication No. US20040092446A1
/ GENERAL INFORMATION:
/ APPLICANT: Goulmy, Els A.J.M.
/ APPLICANT: Hunt, Donald F.
/ APPLICANT: Engelhard, Victor H.
/ TITLE OF INVENTION: HA-1 epitopes and uses thereof
/ FILE REFERENCE: 2183-6047US
/ CURRENT APPLICATION NUMBER: US/10/623,176
/ CURRENT FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: 09/489,760
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: EP 97202303.0
/ PRIOR FILING DATE: 1997-07-23
/ PRIOR APPLICATION NUMBER: PCT/NL98/00424
/ PRIOR FILING DATE: 1998-07-23
/ PRIOR APPLICATION NUMBER: JP 2000-504165
/ PRIOR FILING DATE: 2000-01-24
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 47
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)..(9)
US-10-623-176-47
```

```
Query Match      76.3%; Score 29; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 DDILLEA 9
        | | | | |
Db      2 DDILLEA 7
```

```
RESULT 12
US-10-623-176-48
/ Sequence 48, Application US/10623176
/ Publication No. US20040092446A1
/ GENERAL INFORMATION:
/ APPLICANT: Goulmy, Els A.J.M.
/ APPLICANT: Hunt, Donald F.
/ APPLICANT: Engelhard, Victor H.
/ TITLE OF INVENTION: HA-1 epitopes and uses thereof
/ FILE REFERENCE: 2183-6047US
/ CURRENT APPLICATION NUMBER: US/10/623,176
/ CURRENT FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: 09/489,760
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: EP 97202303.0
/ PRIOR FILING DATE: 1997-07-23
/ PRIOR APPLICATION NUMBER: PCT/NL98/00424
/ PRIOR FILING DATE: 1998-07-23
/ PRIOR APPLICATION NUMBER: JP 2000-504165
/ PRIOR FILING DATE: 2000-01-24
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 48
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
```

```
QY      4 DDILLEA 9
        | | | | |
```

FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-48

Query Match
Best Local Similarity 76.3%; Score 29; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9
Db 2 DDLLEA 7

RESULT 13
US-10-623-176-23
Sequence 23, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 23
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-23

Query Match
Best Local Similarity 71.1%; Score 27; DB 15; Length 9;
Best Local Similarity 85.7%; Pred. No. 1.5e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7
Db 3 VLXDDL 9

RESULT 14
US-10-623-176-40
Sequence 40, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424

PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 40
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-40

Query Match
Best Local Similarity 71.1%; Score 27; DB 15; Length 9;
Best Local Similarity 85.7%; Pred. No. 1.5e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7
Db 3 VLXDDL 9

RESULT 15
US-10-623-176-15
Sequence 15, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 15
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(8)
US-10-623-176-15

Query Match
Best Local Similarity 60.5%; Score 23; DB 15; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9
Db 4 DLLEA 8

Search completed: December 29, 2004, 23:45:47
Job time : 162 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:22:17 ; Search time 38 Seconds
(without alignments)
22.788 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	42.1	7	2	S68004
2	16	42.1	8	2	PC4131
3	14	36.8	7	2	S20446
4	14	36.8	7	2	A59489
5	13	34.2	8	2	A61328
6	13	34.2	9	2	PH0942
7	13	34.2	9	2	A60427
8	12	31.6	4	2	I40697
9	12	31.6	5	2	PT0679
10	12	31.6	5	2	PT0601
11	12	31.6	6	2	B35640
12	12	31.6	6	2	PT0533
13	12	31.6	7	2	A34026
14	12	31.6	7	2	B39040
15	12	31.6	7	2	PT0628
16	12	31.6	7	2	PT0722
17	12	31.6	7	2	PT0576
18	12	31.6	8	2	PT0557
19	12	31.6	9	2	PH0108
20	12	31.6	9	2	PT0562
21	12	31.6	9	2	B30572
22	12	31.6	9	2	PS0253
23	11	28.9	6	2	T11779
24	11	28.9	7	2	S25266
25	11	28.9	8	2	S23428
26	11	28.9	8	2	B33099
27	11	28.9	8	2	S63165
28	11	28.9	8	2	PM0002
29	11	28.9	9	2	S66419

30	11	28.9	9	2	A12872	transaldolase (EC
31	11	28.9	9	2	A61386	macrophage inhibit
32	10	26.3	4	2	A48360	gamma subunit of p
33	10	26.3	4	2	A26209	protein-glutamine
34	10	26.3	6	2	H48394	glycoprotein compo
35	10	26.3	6	2	I48126	alpha-tubulin - Ch
36	10	26.3	7	2	S78024	ribosomal protein
37	10	26.3	7	2	A12016	formylglycinamide
38	10	26.3	7	2	C56793	platelet glycoprot
39	10	26.3	7	2	S29735	polyphosphate-gluc
40	10	26.3	7	2	PC2370	probable H+-transp
41	10	26.3	7	2	T09512	NADH2 dehydrogena
42	10	26.3	8	2	A32523	peptidyl-dipectida
43	10	26.3	8	2	T10077	hypothetical prote
44	10	26.3	8	2	PT0298	Ig heavy chain CRD
45	10	26.3	8	2	A42057	fibroblast growth

ALIGNMENTS

RESULT 1
S68004
hucolin, 75K chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S68004
R:Edgar, P.F.
FEBS Lett. 375, 159-161, 1995
A>Title: Hucolin, a new corticosteroid-binding protein from human plasma with structural
A:Reference number: S68004; MUID:96087107; PMID:7498469
A:Accession: S68004
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <EDG>

Query Match 42.1% Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
DB 4 DDL 6

RESULT 2
PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C:Species: Pseudomonas aeruginosa
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: PC4131
R:Kawasaki, S.; Arai, H.; Igatah, Y.; Kodama, T.
Gene 167, 87-91, 1995
A>Title: Sequencing and characterization of the downstream region of the genes encoding
Y for biosynthesis of heme d1.
A:Reference number: JC4552; MUID:96144254; PMID:8566817
A:Accession: PC4131
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8 <KAW>
A:Cross-references: UNIPROT:P5412; DDBJ:D50473; NID:g1217594
A>Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 42.1% Score 16; DB 2; Length 8;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
DB 2 DDL 4

RESULT 3

S20446
elastase - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C:Accession: S20446
R:Kessler, B.; Sattin, M.; Peretz, M.; Burshtein, Y.
FBS Lett. 299, 291-293, 1992
A>Title: Identification of cleavage sites involved in proteolytic processing of Pseudom
A:Reference number: S20446; MUID:92183956; PMID:1544509
A:Accession: S20446
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <RES>

Query Match
Best Local Similarity 36.8%; Score 14; DB 2; Length 7;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 LLE 8
|||
Db 3 DLID 6

RESULT 4
A59489
protein kinase C inhibitor - rat (fragment)
C:Species: Rattus norvegicus
C:Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 25-Aug-2003
C:Accession: A59489
R:Negoro, M.
Submitted to the Protein Sequence Database, June 2003
A:Description: Purification of PKCI from rat liver.
A:Reference number: A59489
A:Accession: A59489
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <NEG>
A:Experimental source: strain Wistar, liver
A:Note: p-Hydroxyacetophenone-Sepharose binding protein

Query Match
Best Local Similarity 36.8%; Score 14; DB 2; Length 7;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXD 5
|||
Db 2 IFEDD 6

RESULT 5
A61328
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)
C:Species: Balanoptera acutorostrata (minke whale, lesser rorqual)
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 05-Jul-2004
C:Accession: A61328
R:Brictoux-Gregoire, S.; Schyns, R.; Florkin, M.; Emmens, M.; Welling, G.W.; Beintema, J
Biochim. Biophys. Acta 386, 244-255, 1975
A>Title: N-Terminal amino acid sequence of trypsinogen from the lesser rorqual, Balanop
creas.
A:Reference number: A61328; MUID:75146765; PMID:1125273
A:Accession: A61328
A:Molecule type: protein
A:Residues: 1-8 <BRI>
A:Cross-references: UNIPROT:O7M390
C:Keywords: hydrolase; protein digestion; serine protease; zymogen
F:1-8/Domain: activation peptide #status experimental <APT>

Query Match
Best Local Similarity 34.2%; Score 13; DB 2; Length 8;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXKD 5
|||

Db 3 IDDD 6

RESULT 6
PH0942
T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: PH0942
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A>Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A:Reference number: PH0891; MUID:92078857; PMID:1836012
A:Accession: PH0942
A:Molecule type: mRNA
A:Residues: 1-9 <GOL>
A:Experimental source: complete Freund's adjuvant-immunized lymph node
A:Note: the authors translated the codon TGC for residue 2 as Ala
C:Keywords: T-cell receptor

Query Match
Best Local Similarity 34.2%; Score 13; DB 2; Length 9;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLE 8
|||
Db 5 LLE 7

RESULT 7
A60427
macrophage cytotoxicity-inducing factor, 29K - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C:Accession: A60427
R:Jones, C.W.; Prince, C.A.; Williams, J.S.
Exp. Hematol. 19, 704-709, 1991
A>Title: Purification and amino acid analysis of a human macrophage cytotoxicity-inducin
A:Reference number: A60427; MUID:91372335; PMID:1909970
A:Accession: A60427
A:Molecule type: protein
A:Residues: 1-9 <JON>
A:Cross-references: UNIPROT:O7MAR5
A:Note: the sequence from the text on page 706 is inconsistent with that from page 708 i
C:Keywords: cytokine

Query Match
Best Local Similarity 34.2%; Score 13; DB 2; Length 9;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXD 4
|||
Db 4 VLKD 7

RESULT 8
I10697
biotin A - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I10697
R:Shuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A>Title: Transcriptional regulation and gene arrangement of *Bacterichia coli*, *Citrobacte*
A:Reference number: I10697; MUID:89006580; PMID:2971595
A:Accession: I10697
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-4 <RES>
A:Cross-references: UNIPROT:P13071; GB:M21922; NID:G144434

Query Match
Best Local Similarity 31.6%; Score 12; DB 2; Length 4;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
||
Db 3 DD 4

RESULT 9

T-cell receptor beta chain V-D-J region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0679; PT0708

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; PMID:91277601; PMID:1711558

A:Accession: PT0679

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FE2>

A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-2J

A:Accession: PT0708

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FE2>

A:Experimental source: newborn thymus, strain BALB/c, 161-2B

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 31.6%; Score 12; DB 2; Length 5;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
||
Db 4 DD 5

RESULT 10

T-cell receptor beta chain V-D-J region (120-2K) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004

C:Accession: PT0601; PT0617; PT0694

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; PMID:91277601; PMID:1711558

A:Accession: PT0601

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FE2>

A:Cross-references: UNIPROT:Q64512

A:Experimental source: newborn thymus, strain BALB/c, clone 120-2K

A:Accession: PT0617

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FE3>

A:Experimental source: newborn thymus, strain BALB/c, 120-2CA

A:Accession: PT0694

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FE2>

A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1H

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 31.6%; Score 12; DB 2; Length 5;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
||
Db 4 DD 5

RESULT 11

cerebellar degeneration-related protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993

C:Accession: B35640

R:Chen, Y.T.; Retig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.B.

Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990

A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker

A:Reference number: A35640; PMID:90222173; PMID:2326268

A:Accession: B35640

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-6 <CHE>

Query Match

Best Local Similarity 31.6%; Score 12; DB 2; Length 6;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
: ||
Db 3 EDL 5

RESULT 12

T-cell receptor beta chain V-D-J region (126-1AA) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0533

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; PMID:91277601; PMID:1711558

A:Accession: PT0533

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-6 <FE2>

A:Experimental source: adult thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 31.6%; Score 12; DB 2; Length 6;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
||
Db 4 DD 5

RESULT 13

acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)

C:Species: Torpedo californica (Pacific electric ray)

C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 08-Nov-1996

C:Accession: A34026

R:Gidney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedrick, T.; Low, M.G.; Taylor, S.S.; 1

J. Biol. Chem. 263, 1140-1145, 1988

A:Title: Divergence in primary structure between the molecular forms of acetylcholinesterase

A:Reference number: A34026; PMID:88087239; PMID:3335534

A:Accession: A34026

A:Molecule type: protein

A:Residues: 1-7 <GIB>

C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match

Best Local Similarity 31.6%; Score 12; DB 2; Length 7;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 LLEA 9
|| |

Db 1 LIMA 4

RESULT 14

B39040
calsequestrin, fast skeletal muscle - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 23-Feb-1997
C:Accession: B39040
R:Gala, S.E.; Jones, L.R.
J. Biol. Chem. 266, 391-398, 1991
A:Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by casein
A:Reference number: A39040; MUID:91093153; PMID:1985907
A:Accession: B39040
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <CAL>
C:Keywords: phosphoprotein; skeletal muscle

Query Match 31.6%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5

Db 1 DD 2

RESULT 15

PT0628
T-cell receptor beta chain V-D-J region (111-1AG) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0628
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0628
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-7 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 31.6%; Score 12; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5

Db 4 DD 5

Search completed: December 29, 2004, 23:33:55
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:14:21 ; Search time 186 Seconds
(without alignments)
27.841 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLRA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1598

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	42.1	9	Q70SM2	Q70SM2 homo sapien
2	16	42.1	9	Q8LPT5	Q8LPT5 zea mays (m
3	16	42.1	9	P82568	P82568 streptococc
4	16	42.1	9	CAD79467	CAD79467 homo sapi
5	15	39.5	9	Q9XJN0	Q9XJN0 bacterioph
6	14	36.8	9	Q9FSZ2	Q9FSZ2 cicler arlet
7	14	36.8	9	Q9QZAB	Q9QZAB mus musculu
8	14	36.8	9	Q85710	Q85710 rous sarcom
9	14	36.8	9	Q8UTD7	Q8UTD7 human immun
10	13	34.2	7	Q99182	Q99182 gnatholebia
11	13	34.2	8	Q6U7R2	Q6U7R2 cryptococcu
12	13	34.2	8	Q9HDS4	Q9HDS4 aspergillus
13	13	34.2	8	Q7M390	Q7M390 balaenopter
14	13	34.2	8	Q8H9X1	Q8H9X1 bacterioph
15	13	34.2	8	Q8SBJ0	Q8SBJ0 bacterioph
16	13	34.2	8	Q63C68	Q63C68 glycine max
17	13	34.2	8	P72279	P72279 rhodococcus
18	13	34.2	8	Q51594	Q51594 plasmid col
19	13	34.2	8	AAQ88127	AAQ88127 cryptococ
20	13	34.2	8	AAAT01227	AAAT01227 glycine m
21	13	34.2	9	Q7M4R5	Q7M4R5 homo sapien
22	13	34.2	9	Q9B0T4	Q9B0T4 homo sapien
23	13	34.2	9	Q6Q7G0	Q6Q7G0 rana ridib
24	13	34.2	9	Q6UVK2	Q6UVK2 malus domes
25	13	34.2	9	Q91IX6	Q91IX6 macaca neme
26	13	34.2	9	Q8AUM7	Q8AUM7 carassius a
27	13	34.2	9	AAQ57275	AAQ57275 malus dom
28	13	34.2	9	AAAT01544	AAAT01544 rana ridi
29	13	34.2	9	AAAT01545	AAAT01545 rana ridi
30	13	34.2	9	AAAT01546	AAAT01546 rana ridi
31	13	34.2	9	AAAT01547	AAAT01547 rana ridi

32	13	34.2	9	AAAT01548	AAAT01548 rana ridi
33	13	34.2	9	AAAT01549	AAAT01549 rana ridi
34	13	34.2	9	AAAT01550	AAAT01550 rana ridi
35	13	34.2	9	AAAT01551	AAAT01551 rana ridi
36	13	34.2	9	AAAT01552	AAAT01552 rana ridi
37	13	34.2	9	AAAT01553	AAAT01553 rana ridi
38	13	34.2	9	AAAT01554	AAAT01554 rana ridi
39	13	34.2	9	AAAT01555	AAAT01555 rana ridi
40	13	34.2	9	AAAT01556	AAAT01556 rana ridi
41	13	34.2	9	AAAT01557	AAAT01557 rana ridi
42	13	34.2	9	AAAT01558	AAAT01558 rana ridi
43	13	34.2	9	AAAT01559	AAAT01559 rana ridi
44	13	34.2	9	AAAT01560	AAAT01560 rana ridi
45	12	31.6	5	BIOA_CITFR	P13071 Citrobacter

ALIGNMENTS

RESULT 1	Q70SM2	PRELIMINARY;	PRT;	9 AA.
ID	Q70SM2			
AC	Q70SM2			
DT	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Hypermethylation in cancer 1 (Fragment).			
GN	Name=HIC1,			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Pine S., Guerardel C., Deltour S., Godwin A.K., Lepince D.;			
RT	"Identification of a second G-C-rich promoter conserved in the human,			
RT	murine and rat tumor suppressor genes HIC1.";			
DR	Oncogene 23:4023-4031 (2004).			
DR	EMBL: AJ550615; CAD79467.1; --			
FT	NON_TER			
SQ	SEQUENCE 9 AA; 964 MW; 5B56DBD1681AAA7 CRC64;			
Query Match	42.1%; Score 16; DB 2; Length 9;			
Best Local Similarity	60.0%; Pred. No. 1.8e+06;			
Matches	3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Qy	5 DLEA 9			
Db	3 DTMEA 7			
RESULT 2	Q8LPT5	PRELIMINARY;	PRT;	9 AA.
ID	Q8LPT5			
AC	Q8LPT5			
DT	01-OCT-2002 (TrEMBLrel. 22, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	Beta-expansin-like protein (Fragment).			
OS	Zea mays (Maize).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Poaceae; Panicoideae; Andropogoneae; Zea.			
NCBI_Taxid=4577;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,			
RA	Morgan M., Rafalski J.A.;			
DT	Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY094310; AAM21836.1; --			
FT	NON_TER			
SQ	SEQUENCE 9 AA; 977 MW; 5C05BD2DCB1AAA3 CRC64;			

Query Match 42.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 33.3%; Pred. No. 1.8e+06;
 Matches 2; Conservative 4; Mismatches 0; Indels 0;

QY 4 DDLEA 9
 ||::||
 DB 4 DEVDA 9

RESULT 3

PRELIMINARY; PRT; 9 AA.
 ID P82568
 AC P82568;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Unknown protein from 2D-page (Fragment).
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC STRAIN=CRS4;
 RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
 RA Vanbogaellen R.A.;
 RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
 RT proteins";
 RL Submitted (MAY-2000) to Swiss-Prot.
 CC -1- MASS SPECTROMETRY; MW=22592.04; METHOD=Electrospray.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 9 AA; 1069 MW; 2A771042CB1AB2D7 CRC64;

Query Match 42.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 3; Mismatches 0; Indels 0;

QY 4 DDLE 8
 ||::||
 DB 4 DEVIE 8

RESULT 4

PRELIMINARY; PRT; 9 AA.
 ID CAD79467
 AC CAD79467;
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypermethylated in cancer 1 (Fragment).
 GN HIC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX PubMed=15007385;
 RA Pinte S., Guevariel C., Delbour S., Godwin A.K., Leprience D.,
 RT "Identification of a second G-C-rich promoter conserved in the human,
 RT murine and rat tumor suppressor genes HIC1";
 RL Oncogene 23:4023-4031 (2004).
 DR EMBL; AJ550616; CAD79467.1; -.
 FT NON TER 9
 FT NON TER 9
 SQ SEQUENCE 9 AA; 964 MW; 5B5E6DB1681AA7 CRC64;

Query Match 42.1%; Score 16; DB 2; Length 9;
 Best Local Similarity 60.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 DLEA 9
 ||::||
 DB 3 DTMEA 7

RESULT 5

PRELIMINARY; PRT; 9 AA.
 ID O9XJN0
 AC O9XJN0;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE P10 (Fragment).
 OS Bacteriophage phi-10.
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=90889;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99350412; PubMed=10419946;
 RA Mindich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,
 RA Hoogstraaten D.;
 RT "Isolation of additional bacteriophages with genomes of segmented
 RT double-stranded RNA";
 RL J. Bacteriol. 181:4505-4508 (1999).
 DR EMBL; AF125675; AAD22555.1; -.
 FT NON TER 9
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1058 MW; 880E376AA720544A CRC64;

Query Match 39.5%; Score 15; DB 2; Length 9;
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 3; Mismatches 0; Indels 0;

QY 4 DDLE 8
 ||::||
 DB 2 DNILD 6

RESULT 6

PRELIMINARY; PRT; 9 AA.
 ID O9FSZ2
 AC O9FSZ2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Etisolated epicotyls;
 RA Dopico B., Jimenez T., Labrador E.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ299069; CAC10216.1; -.
 KW Hypothetical protein.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 9 AA; 990 MW; 9441BDDA7272EBE CRC64;

Query Match 36.8%; Score 14; DB 2; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLEA 9
 ||::||
 DB 4 LLDA 7

RESULT 7

PRELIMINARY; PRT; 9 AA.
 ID O9QZAB
 AC O9QZAB;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE P10 (Fragment).
 OS Bacteriophage phi-10.
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=90889;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99350412; PubMed=10419946;
 RA Mindich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,
 RA Hoogstraaten D.;
 RT "Isolation of additional bacteriophages with genomes of segmented
 RT double-stranded RNA";
 RL J. Bacteriol. 181:4505-4508 (1999).
 DR EMBL; AF125675; AAD22555.1; -.
 FT NON TER 9
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1058 MW; 880E376AA720544A CRC64;

```

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE C-type lectin DC1A (Fragment).
GN Name=DC11;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gorzki K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tauchiya H.;
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF192526; AAF04843.1; -.
DR MGD; MGI:2136650; Dcl1.
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR GO; GO:0046703; F:natural killer cell lectin-like receptor bi. .; IDA.
DR GO; GO:0005515; F:protein binding; IPT.
KW Lectin.
FT NON TER
SQ SEQUENCE 9 AA; 994 MW; 342161AB172EBAB7 CRC64;

Query Match
Best Local Similarity 36.8%; Score 14; DB 2; Length 9;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DLE 8
DB 3 DCLE 6

RESULT 8
ID 085710 PRELIMINARY; PRT; 9 AA.
AC 085710;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Pol protein (Fragment).
OS Rous sarcoma virus.
OC Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxId=11886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84115080; PubMed=6319754;
RA Lerner T.L., Hanafusa H.;
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:
RT extent of env deletion and possible genealogical relationship with
RT other viral strains."
RL J. Virol. 49:549-556(1984).
DR EMBL; K03365; AAA42557.1; -.
FT NON TER
SQ SEQUENCE 9 AA; 949 MW; 94AA144DDDD731AA CRC64;

Query Match
Best Local Similarity 36.8%; Score 14; DB 2; Length 9;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DDL 9
DB 1 EDT 6

RESULT 9
ID 08UTD7 PRELIMINARY; PRT; 9 AA.
AC 08UTD7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Vpu protein.
GN Name=vpu;
OS Human immunodeficiency virus 1.

```

```

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21988475; PubMed=11991972;
RA Novitsky V., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,
RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,
RA Foley B.T., Gaoletke S., Rybak N., Gaselstwe S., Vanberg F.,
RA Marink R., Lee T.-H., Essex M.;
RT "Human immunodeficiency virus type 1 subtype C molecular phylogeny:
RT consensus sequence for an AIDS vaccine design?";
RL J. Virol. 76:5435-5451(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Novitsky V.A., McLane M.F., Chigwedere P., Ndung'u T., Klein I.,
RA Chang S.-Y., Peter T., Thior I., Rybak N., Gaselstwe S., Vanberg F.,
RA Marink R., Lee T.-H., Essex M.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF443091; AAL34712.1; -.
SQ SEQUENCE 9 AA; 1102 MW; 188BD40B17272440 CRC64;

Query Match
Best Local Similarity 36.8%; Score 14; DB 2; Length 9;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLE 8
DB 3 NLE 6

RESULT 10
ID 099182 PRELIMINARY; PRT; 7 AA.
AC 099182;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome oxidase I (Fragment).
GN Name=COI;
OS Gnathobias zonatus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnathobias.
OX NCBI_TaxId=135316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20072928; PubMed=10603257;
RA Murphy W.J., Thomsen J.E., Collier G.E.;
RT "Phylogeny of the Neotropical Killifish family Rivulidae
RT (Cyprinodontiformes, Aplocheilidae) inferred from mitochondrial DNA
RT sequences."
RL Mol. Phylogenet. Evol. 13:289-301(1999).
DR EMBL; AF002591; AAD01074.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 7 AA; 899 MW; 672721F6C8572030 CRC64;

Query Match
Best Local Similarity 34.2%; Score 13; DB 2; Length 7;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLX 7
DB 1 ILX 7

RESULT 11
ID 06UTR2 PRELIMINARY; PRT; 8 AA.
AC 06UTR2;

```

DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Squalene monooxygenase (EC 1.14.99.7) (Fragment).
 GN Name=ERG1.
 OS Cryptococcus neoformans var. grubii H99.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=235443;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H99;
 RA Stuart L.T., Allen A., Dietrich F.S.;
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY376718; AA088127.1; -.
 DR GO; GO:0004506; F:squalene monooxygenase activity; IEA.
 KW Monooxygenase; Oxidoreductase.
 FT NON TER 1 1
 SQ SEQUENCE 8 AA; 811 MW; 0467776AAB1D727 CRC64;

Query Match 34.2%; Score 13; DB 2; Length 8;
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDD 5
 :|:
 Db 1 LLAED 5

RESULT 12

ID Q9HDS4 PRELIMINARY; PRT; 8 AA.
 AC Q9HDS4;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE TPC polypeptide (Fragment).
 GN Name=tpc;
 OS Aspergillus flavus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5059;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A55;
 RC MEDLINE=21173963; PubMed=11273679;
 RA Geisler D.M., Dörner J.W., Horn B.W., Taylor J.W.;
 RT "The phylogenetics of mycotoxin and sclerotium production in
 RT Aspergillus flavus and Aspergillus oryzae.";
 RL Fungal Genet. Biol. 31:169-179(2000).
 DR EMBL; AF261861; AAG16135.1; -.
 KW Polypeptide.
 FT NON TER 8 8
 SQ SEQUENCE 8 AA; 807 MW; F3B2C72AB5B87DD6 CRC64;

Query Match 34.2%; Score 13; DB 2; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLE 8
 ||:
 Db 5 DLVD 8

RESULT 13

ID Q7M390 PRELIMINARY; PRT; 8 AA.
 AC Q7M390;
 DT 01-MAR-2004 (TREMBlrel. 26, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
 DE Trypsin (EC 3.4.21.4) precursor (Fragment).
 OS Balaenoptera acutorostrata (Minke whale) (Lesser rorqual).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
 OC Balaenopteridae; Balaenoptera.
 OX NCBI_TaxID=9767;
 RN [1]
 RP SEQUENCE.

RA Briteux-Gregoire S., Schyns R., Florin M., Emmens M., Welling G.W.,
 RA Beintema J.J.;
 RT "N-terminal amino acid sequence of trypsinogen from the lesser
 RT rorqual, Balaenoptera acutorostrata (Cetacea). Simultaneous isolation
 RT of trypsinogen, chymotrypsinogen and ribonuclease from pancreas.";
 RL Biochim. Biophys. Acta 386:244-255(1975).
 DR PIR; A61328; A61328.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 FT NON TER 1 1
 SQ SEQUENCE 8 AA; 964 MW; 8D3A8AAAAA04769 CRC64;

Query Match 34.2%; Score 13; DB 2; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXXD 5
 :|:
 Db 3 IDDD 6

RESULT 14

ID Q8H9X1 PRELIMINARY; PRT; 8 AA.
 AC Q8H9X1;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Gp30.2 protein (Fragment).
 GN Name=30.2;
 OS Bacteriophage L210.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=192973;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kolesnikiene G., Nivinskas R.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ458400; CAD30256.1; -.
 FT NON TER 8 8
 SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 34.2%; Score 13; DB 2; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXD 4
 :|:
 Db 5 ILTD 8

RESULT 15

ID Q8SBJ0 PRELIMINARY; PRT; 8 AA.
 AC Q8SBJ0;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Gp30.2 (Fragment).
 GN Name=30.2;
 OS Bacteriophage RB69.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=12353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Plesniene L., Kolesnikiene G., Truncate L., Zajackauskaite A.,

RA Nivinskas R.;
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ439452; CAD28423.1; -
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 34.2%; Score 13; DB 2; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXD 4
 : | |
 Db 5 ILTD 8

Search completed: December 29, 2004, 23:33:08
 Job time : 191 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:57:00 ; Search time 151 Seconds
(without alignments)
21.381 Million cell updates/sec

Title: US-10-791-217A-1

Perfect score: 38
Sequence: 1 VLXDDLLEA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	9	2	AAW99195
2	36	94.7	9	2	AAW99196
3	36	94.7	9	2	AAW99197
4	36	94.7	9	2	AAW97572
5	36	94.7	9	2	AAW97375
6	36	94.7	9	2	AAW97374
7	36	94.7	9	8	ADH40333
8	36	94.7	9	8	ADH40334
9	36	94.7	13	2	AAW99199
10	36	94.7	13	2	AAW99198
11	36	94.7	13	2	AAW97415
12	36	94.7	13	2	AAW97414
13	36	94.7	1136	5	AD117193
14	36	94.7	1136	5	AD116744
15	36	94.7	1136	8	ADH61299
16	36	94.7	1136	8	ADH42338
17	36	94.7	1165	5	AD117194
18	36	94.7	192	2	AAV34986
19	36	94.7	226	6	AAU53226
20	36	94.7	226	6	ABM49745
21	36	94.7	499	5	ABR90575
22	36	94.7	499	6	ABR26874
23	36	94.7	635	7	ABO82267
24	36	94.7	4150	3	AAV92707
25	36	94.7	149	3	AAH25210

26	31	81.6	276	5	ABB48245	Abd48245	Listeria
27	31	81.6	289	5	AAU72987	Aau72987	Neisseria
28	31	81.6	289	5	AAU73000	Aau73000	Neisseria
29	31	81.6	289	5	ABG91578	Abg91578	Purine/py
30	31	81.6	289	5	ABG91584	Abg91584	Purine/py
31	31	81.6	289	6	ABP77961	Abp77961	N. gonorr
32	31	81.6	289	6	ABU37265	Abu37265	Protein e
33	31	81.6	289	6	ABU37881	Abu37881	Protein e
34	31	81.6	665	6	ABU31214	Abu31214	Protein e
35	31	81.6	675	6	ABU28302	Abu28302	Protein e
36	31	81.6	1042	4	ABG59345	Abg59345	Drosophila
37	31	81.6	1177	6	ABP71254	Abp71254	Chimeric
38	31	81.6	1186	6	ABP71253	Abp71253	Zebrafish
39	30	78.9	94	4	AAU86478	Aau86478	Novel hum
40	30	78.9	94	7	ADB59812	Adb59812	Connectiv
41	30	78.9	117	8	ADJ48890	Adj48890	Oil-assoc
42	30	78.9	162	8	ADJ20159	Adj20159	Human PRO
43	30	78.9	328	4	AAU73861	Aau73861	Human COL
44	30	78.9	423	6	ABU21143	Abu21143	Protein e
45	30	78.9	454	7	ABO64879	AbO64879	Klebsiella

ALIGNMENTS

RESULT 1	AAW99195	standard; peptide; 9 AA.
ID	AAW99195	
AC	AAW99195;	
DT	20-MAY-1999	(first entry)
XX		
DE	Minor histocompatibility antigen HA-1 T-cell epitope #1.	
XX		
KW	Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;	
KM	graft versus host disease; bone marrow transplant; leukaemia; vaccine;	
KW	diagnosis; aplastic anaemia; immune deficiency disease.	
OS	Homo sapiens.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	Misc-difference 3	/label= His, Arg
XX		
PN	WO9905174-A1.	
XX		
PD	04-FEB-1999.	
XX		
XX	23-JUL-1998;	98WO-NL000425.
PR	23-JUL-1997;	97BP-00202303.
XX		
XX	(UYLE-) RIKXSUNIV LEIDEN.	
XX		
PI	Goulmy EAMV, Hunt DR, Engelhard VH;	
XX		
DR	WPI, 1999-15312/13.	
XX		
PT	A new minor histocompatibility antigen, HA-1 - useful to treat immune	
PT	diseases and prevent rejection and host versus graft disease in bone	
PT	marrow and organ transplantation.	
XX		
PS	Claim 1; Page 32; 47pp; English.	
XX		
CC	The present sequence represents a new peptide (P1) constituting a T-cell	
CC	epitope obtainable from the minor histocompatibility antigen HA-1. The	
CC	peptide is immunogenic and can be used as part of a vaccine. P1 is used	
CC	as a medicine, to induce tolerance for transplants, prevent rejection	
CC	and/or graft versus host disease, or to treat (auto) immune diseases. In	
CC	particular it can be used with bone marrow transplantation, in the	
CC	treatment of severe aplastic anaemia, leukaemia, and immune deficiency	

CC diseases
 XX
 SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9
 |||||
 Db 1 VLXDDLLEA 9

RESULT 2
 AAM99196
 ID AAM99196 standard; peptide; 9 AA.
 AC AAM99196;
 XX
 DT 20-MAY-1999 (first entry)
 XX
 DE Minor histocompatibility antigen HA-1 T-cell epitope #2.
 XX
 KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
 KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
 KW diagnosis; aplastic anaemia; immune deficiency disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9905174-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 23-JUL-1998; 98WO-NL000425.
 XX
 PR 23-JUL-1997; 97EP-00202303.
 XX
 PA (UYLE-) RIKKSUNIV LEIDEN.
 XX
 PI Goulmy EAJM, Hunt DF, Engelhard VH;
 XX
 DR WPI; 1999-153312/13.
 XX
 PT A new minor histocompatibility antigen, HA-1 - useful to treat immune
 PT diseases and prevent rejection and host versus graft disease in bone
 PT marrow and organ transplantation.
 XX
 PS Claim 3; Page 32; 47pp; English.
 XX
 CC The present sequence represents a new peptide (P1) constituting a T-cell
 CC epitope obtainable from the minor histocompatibility antigen HA-1. The
 CC peptide is immunogenic and can be used as part of a vaccine. P1 is used
 CC as a medicine, to induce tolerance for transplants, prevent rejection
 CC and/or graft versus host disease, or to treat (auto) immune diseases. In
 CC particular it can be used with bone marrow transplantation, in the
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency
 CC diseases
 CC
 SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
 Best Local Similarity 88.9%; Pred. No. 1.7e+06;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9
 |||||
 Db 1 VLXDDLLEA 9

RESULT 3
 AAM99197
 ID AAM99197 standard; peptide; 9 AA.
 XX

AC AAM99197;
 XX
 DT 20-MAY-1999 (first entry)
 XX
 DE Minor histocompatibility antigen HA-1 T-cell epitope #3.
 XX
 KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
 KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
 KW diagnosis; aplastic anaemia; immune deficiency disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9905174-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 23-JUL-1998; 98WO-NL000425.
 XX
 PR 23-JUL-1997; 97EP-00202303.
 XX
 PA (UYLE-) RIKKSUNIV LEIDEN.
 XX
 PI Goulmy EAJM, Hunt DF, Engelhard VH;
 XX
 DR WPI; 1999-153312/13.
 XX
 PT A new minor histocompatibility antigen, HA-1 - useful to treat immune
 PT diseases and prevent rejection and host versus graft disease in bone
 PT marrow and organ transplantation.
 XX
 PS Disclosure; Page 15; 47pp; English.
 XX
 CC The present sequence represents a new peptide (P1) constituting a T-cell
 CC epitope obtainable from the minor histocompatibility antigen HA-1. The
 CC peptide is immunogenic and can be used as part of a vaccine. P1 is used
 CC as a medicine, to induce tolerance for transplants, prevent rejection
 CC and/or graft versus host disease, or to treat (auto) immune diseases. In
 CC particular it can be used with bone marrow transplantation, in the
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency
 CC diseases
 CC
 SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
 Best Local Similarity 88.9%; Pred. No. 1.7e+06;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9
 |||||
 Db 1 VLXDDLLEA 9

RESULT 4
 AAM97572
 ID AAM97572 standard; peptide; 9 AA.
 AC AAM97572;
 XX
 DT 20-MAY-1999 (first entry)
 XX
 DE T-cell epitope from the minor histocompatibility antigen HA-1.
 XX
 KW T-cell epitope; minor histocompatibility antigen HA-1; vaccine;
 KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;
 KW neoplastic haematopoietic cell.
 XX
 OS Homo sapiens.
 XX
 PN WO9905173-A1.

Key Location/Qualifiers
 FH Misc-difference 3 /note= "His or Arg"
 FT
 XX
 PN

XX 04-FEB-1999.
 PD 23-JUL-1998; 98WO-NL000424.
 XX 23-JUL-1997; 97EP-00202303.
 PR 23-JUL-1997; 97EP-00202303.
 XX (UYLE-) RIJXSUNIV LEIDEN.
 PA
 XX Goulmy EAM, Hunt DF, Engelhard VH;
 PI WPI; 1999-142855/12.
 DR
 XX Immunogenic peptide from minor histocompatibility antigen HA-1 - useful
 PT for inducing tolerance to transplants and prevent rejection or graft-
 PT versus-host disease.
 XX
 PS Claim 1; Page 39; 57pp; English.
 CC The present sequence represents an immunogenic peptide constituting a T-
 CC cell epitope, obtainable from the minor histocompatibility antigen HA-1.
 CC The peptide can be used in vaccines or pharmaceutical formulations as
 CC medicines to induce tolerance for transplants so as to prevent rejection
 CC and/or Graft-versus-Host Disease, or to treat autoimmune diseases.
 CC Neoplastic haematopoietic cells presenting the peptides, in an HLA class
 CC I context, can be eliminated after specific recognition of the peptides.
 CC The peptides can also be used to raise antibodies, T-cell receptor, B-
 CC and T-cells
 XX
 SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
 |||||
 Db 1 VLXDDLLEA 9

RESULT 5
 AAW97375
 ID AAW97375 standard; protein; 9 AA.
 XX
 AC AAW97375;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE HA-1 H-allele sequence.
 XX
 KM Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
 KM R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
 KM severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09905313-A2.
 PD 04-FEB-1999.
 XX
 PF 23-JUL-1998; 98WO-EP004928.
 XX
 PR 23-JUL-1997; 97EP-00202303.
 PR 02-JUN-1998; 98EP-00870125.
 XX
 PA (UYLE-) RIJXSUNIV LEIDEN.
 XX
 PI Goulmy E;
 XX WPI; 1999-142860/12.
 DR
 PT Typing minor histocompatibility antigen HA-1 - by amplifying and
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection

PT of genetic aberrances.
 XX
 PS Claim 18; Fig 5; 59pp; English.
 XX

CC The present sequence represents part of the minor histocompatibility
 CC antigen HA-1 H-allele. The specification describes methods for typing
 CC alleles (preferably the H and R alleles) of the minor histocompatibility
 CC antigen HA-1 in a sample, which comprise detecting polymorphic
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
 CC methods can be used for HA-1 typing for bone marrow transplants, severe
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
 CC detection of genetic aberrances. The probes and primers of the invention
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
 CC anti-idiotypic B cells and/or T cells and antibodies
 XX
 SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
 Best Local Similarity 88.9%; Pred. No. 1.7e+06;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
 |||||
 Db 1 VLXDDLLEA 9

RESULT 6
 AAW97374
 ID AAW97374 standard; protein; 9 AA.
 XX
 AC AAW97374;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE HA-1 R-allele sequence.
 XX
 KM Intron; minor histocompatibility antigen HA-1; typing allele; H allele;
 KM R allele; polymorphic nucleotide; HA-1 typing; bone marrow transplant;
 KM severe aplastic anaemia; leukaemia; immune deficiency disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN W09905313-A2.
 PD 04-FEB-1999.
 XX
 PF 23-JUL-1998; 98WO-EP004928.
 XX
 PR 23-JUL-1997; 97EP-00202303.
 PR 02-JUN-1998; 98EP-00870125.
 XX
 PA (UYLE-) RIJXSUNIV LEIDEN.
 XX
 PI Goulmy E;
 XX WPI; 1999-142860/12.
 DR
 PT Typing minor histocompatibility antigen HA-1 - by amplifying and
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g. detection
 PT of genetic aberrances.
 XX
 PS Claim 13; Fig 5; 59pp; English.
 XX
 CC The present sequence represents part of the minor histocompatibility
 CC antigen HA-1 R-allele. The specification describes methods for typing
 CC alleles (preferably the H and R alleles) of the minor histocompatibility
 CC antigen HA-1 in a sample, which comprise detecting polymorphic
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The
 CC methods can be used for HA-1 typing for bone marrow transplants, severe
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as
 CC detection of genetic aberrances. The probes and primers of the invention
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be used
 CC anti-idiotypic B cells and/or T cells and antibodies

XX
SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 2; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||
1 VLRDDLEA 9

RESULT 7
ADH40333
ID ADH40333 standard; peptide; 9 AA.

XX AC ADH40333;

XX DT 11-MAR-2004 (first entry)

XX DE Human minor histocompatibility antigen HA-1 T cell epitope.

XX KM human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
XX minor histocompatibility antigen; mHag; T cell epitope.

XX OS Homo sapiens.

XX PN WO2003106692-A2.

XX PD 24-DEC-2003.

XX PF 13-JUN-2003; 2003WO-EP006251.

XX PR 13-JUN-2002; 2002EP-00013423.

XX PA (MERCK) MERCK PATENT GMBH.

XX PI Strittmatter W, Moll H;

XX PS Strittmatter W, Moll H;

XX DR WPI; 2004-082200/08.

XX PT Providing allelic variant epitope of protein based on single nucleotide
XX polymorphism by defining target protein, screening database of protein,
XX identifying, selecting allelic variant protein, creating variant
XX epitopes.

XX PS Disclosure; Page 82; 119pp; English.

XX CC The invention relates to a novel method for providing epitopes of allelic
XX variants of antigenic proteins from specific species based on single
XX nucleotide polymorphism (SNP), by defining target protein/peptide or its
XX subset, screening database of DNA encoding target protein, identifying,
XX selecting allelic peptide/protein variants, expression product or its
XX fragment encoded by DNA sequence having SNP, creating variant epitopes,
XX selecting epitopes binding to MHC protein. A protein of the invention has
XX cytostatic activity, and may have a use in a vaccine. The method is
XX useful for generating a SNP profile of one or more individuals from a
XX given species by applying the method for several protein from the
XX individuals, where the SNP profile was related to disease, preferably
XX cancer. This is useful for diagnosing a disease in an individual by
XX generating the SNP-related polymorphic profile. A method of the invention
XX is useful for transplanting haematopoietic stem cells from a donor to a
XX recipient and treating cancer, preferably leukaemia, and for determining
XX the progression, regression or onset of a treated disease. The present
XX sequence is used in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9

DB 1 VLHDDLLEA 9
|||
1 VLRDDLEA 9

RESULT 8
ADH40334
ID ADH40334 standard; peptide; 9 AA.

XX AC ADH40334;

XX DT 11-MAR-2004 (first entry)

XX DE Human minor histocompatibility antigen HA-1 T cell epitope.

XX KM human; cytostatic; vaccine; SNP profile; cancer; leukaemia;
XX minor histocompatibility antigen; mHag; T cell epitope.

XX OS Homo sapiens.

XX PN WO2003106692-A2.

XX PD 24-DEC-2003.

XX PF 13-JUN-2003; 2003WO-EP006251.

XX PR 13-JUN-2002; 2002EP-00013423.

XX PA (MERCK) MERCK PATENT GMBH.

XX PI Strittmatter W, Moll H;

XX PS Strittmatter W, Moll H;

XX DR WPI; 2004-082200/08.

XX PT Providing allelic variant epitope of protein based on single nucleotide
XX polymorphism by defining target protein, screening database of protein,
XX identifying, selecting allelic variant protein, creating variant
XX epitopes.

XX PS Disclosure; Page 82; 119pp; English.

XX CC The invention relates to a novel method for providing epitopes of allelic
XX variants of antigenic proteins from specific species based on single
XX nucleotide polymorphism (SNP), by defining target protein/peptide or its
XX subset, screening database of DNA encoding target protein, identifying,
XX selecting allelic peptide/protein variants, expression product or its
XX fragment encoded by DNA sequence having SNP, creating variant epitopes,
XX selecting epitopes binding to MHC protein. A protein of the invention has
XX cytostatic activity, and may have a use in a vaccine. The method is
XX useful for generating a SNP profile of one or more individuals from a
XX given species by applying the method for several protein from the
XX individuals, where the SNP profile was related to disease, preferably
XX cancer. This is useful for diagnosing a disease in an individual by
XX generating the SNP-related polymorphic profile. A method of the invention
XX is useful for transplanting haematopoietic stem cells from a donor to a
XX recipient and treating cancer, preferably leukaemia, and for determining
XX the progression, regression or onset of a treated disease. The present
XX sequence is used in the exemplification of the invention.

XX SQ Sequence 9 AA;

Query Match 94.7%; Score 36; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||
1 VLRDDLEA 9

RESULT 9
AAW99199
ID AAW99199 standard; peptide; 13 AA.
XX

```

AC AAW99199;
XX
XX 20-MAY-1999 (first entry)
XX
XX VR cell K1A0223 protein sequence.
XX
XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
XX graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX Synthetic.
XX
XX WO9905174-A1.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-NL000425.
XX
XX 23-JUL-1997; 97EP-00202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Goulmy EAJM, Hunt DF, Engelhard VH;
XX
XX WPI; 1999-153312/13.
XX
XX N-PSDB; AAX19409.
XX
XX A new minor histocompatibility antigen, HA-1 - useful to treat immune
XX diseases and prevent rejection and host versus graft disease in bone
XX marrow and organ transplantation.
XX
XX Disclosure; Page 31; 47pp; English.
XX
XX The present invention describes a new peptide (P1) constituting a T-cell
XX epitope obtainable from the minor histocompatibility antigen HA-1. The
XX peptide is immunogenic and can be used as part of a vaccine. P1 is used
XX as a medicine, to induce tolerance for transplants, prevent rejection
XX and/or graft versus host disease, or to treat (auto) immune diseases. In
XX particular it can be used with bone marrow transplantation, in the
XX treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX diseases. The present sequence represents a K1A0223 sequence given in
XX the present invention.
XX
XX Sequence 13 AA;
XX
XX Query Match 94.7%; Score 36; DB 2; Length 13;
XX Best Local Similarity 88.9%; Pred. No. 1.5;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 VLXDDLEA 9
XX |||||
XX 3 VLHDDLEA 11
XX
XX RESULT 10
XX AAW99198
XX ID AAW99198 standard; peptide; 13 AA.
XX
XX AAW99198;
XX
XX 20-MAY-1999 (first entry)
XX
XX DH cell K1A0223 protein sequence.
XX
XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
XX graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX Synthetic.
XX
XX WO9905174-A1.
XX
XX 04-FEB-1999.
XX
XX
XX

```

```

XX
XX 23-JUL-1998; 98WO-NL000425.
XX
XX 23-JUL-1997; 97EP-00202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Goulmy EAJM, Hunt DF, Engelhard VH;
XX
XX WPI; 1999-153312/13.
XX
XX N-PSDB; AAX19408.
XX
XX A new minor histocompatibility antigen, HA-1 - useful to treat immune
XX diseases and prevent rejection and host versus graft disease in bone
XX marrow and organ transplantation.
XX
XX Disclosure; Page 31; 47pp; English.
XX
XX The present invention describes a new peptide (P1) constituting a T-cell
XX epitope obtainable from the minor histocompatibility antigen HA-1. The
XX peptide is immunogenic and can be used as part of a vaccine. P1 is used
XX as a medicine, to induce tolerance for transplants, prevent rejection
XX and/or graft versus host disease, or to treat (auto) immune diseases. In
XX particular it can be used with bone marrow transplantation, in the
XX treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX diseases. The present sequence represents a K1A0223 sequence given in
XX the present invention.
XX
XX Sequence 13 AA;
XX
XX Query Match 94.7%; Score 36; DB 2; Length 13;
XX Best Local Similarity 88.9%; Pred. No. 1.5;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 VLXDDLEA 9
XX |||||
XX 3 VLHDDLEA 11
XX
XX Db
XX
XX RESULT 11
XX AAW97415
XX ID AAW97415 standard; protein; 13 AA.
XX
XX AAW97415;
XX
XX 20-MAY-1999 (first entry)
XX
XX K1A0223 polymorphism in HA-1 positive homozygous individuals.
XX
XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;
XX transplant rejection; Graft-versus-Host Disease; autoimmune disease;
XX neoplastic haematopoietic cell; K1A0223 polymorphism.
XX
XX Homo sapiens.
XX
XX WO9905173-A1.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-NL000424.
XX
XX 23-JUL-1997; 97EP-00202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Goulmy EAJM, Hunt DF, Engelhard VH;
XX
XX WPI; 1999-142855/12.
XX
XX N-PSDB; AAX16081.
XX
XX Immunogenic peptide from minor histocompatibility antigen HA-1 - useful
XX for inducing tolerance to transplants and prevent rejection or graft-
XX versus-host disease.
XX

```

XX PS Disclosure; Page 38; 57pp; English.

CC CC The present sequence represents the KIAA0223 polymorphism in HA-1

CC positive homoygous individuals. The specification describes an

CC immunogenic peptide constituting a T-cell epitope, obtainable from the

CC minor histocompatibility antigen HA-1. The peptide can be used in

CC vaccines or pharmaceutical formulations as medicines to induce tolerance

CC for transplants so as to prevent rejection and/or graft-versus-host

CC disease, or to treat autoimmune diseases. Neoplastic haematopoietic cells

CC presenting the peptides, in an HLA class I context, can be eliminated

CC after specific recognition of the peptides. The peptides can also be used

CC to raise antibodies, T-cell receptor, B- and T-cells

XX SQ Sequence 13 AA;

Query Match 94.7%; Score 36; DB 2; Length 13;

Best Local Similarity 88.9%; Pred. No. 1.5;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9

Db 3 VLHDDLLEA 11

RESULT 12

AAW97414

ID AAW97414 standard; protein; 13 AA.

AC AAW97414;

XX 20-MAY-1999 (first entry)

DT

XX KIAA0223 polymorphism in HA-1 negative homoygous individuals.

DE

XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;

KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;

KW neoplastic haematopoietic cell; KIAA0223 polymorphism.

XX Homo sapiens.

OS W0905173-A1.

PN W0905173-A1.

XX 04-FEB-1999.

PD

XX 23-JUL-1998; 98WO-NL000424.

PF

XX 23-JUL-1997; 97EP-00202303.

PR

XX (UYLE-) RICKSUNITV LEIDEN.

PA

XX Goulmv EAJM, Hunt DF, Engelhard VH;

PI

XX WPI, 1999-142855/12.

DR

XX N-PSDB; AAX16080.

DR

XX Immunogenic peptide from minor histocompatibility antigen HA-1 - useful

PT for inducing tolerance to transplants and prevent rejection or graft-

PT versus-host disease.

XX

XX Disclosure; Page 38; 57pp; English.

PS

XX The present sequence represents the KIAA0223 polymorphism in HA-1

CC negative homoygous individuals. The specification describes an

CC immunogenic peptide constituting a T-cell epitope, obtainable from the

CC minor histocompatibility antigen HA-1. The peptide can be used in

CC vaccines or pharmaceutical formulations as medicines to induce tolerance

CC for transplants so as to prevent rejection and/or Graft-versus-Host

CC disease, or to treat autoimmune diseases. Neoplastic haematopoietic cells

CC presenting the peptides, in an HLA class I context, can be eliminated

CC after specific recognition of the peptides. The peptides can also be used

CC to raise antibodies, T-cell receptor, B- and T-cells

XX

SQ Sequence 13 AA;

Query Match 94.7%; Score 36; DB 2; Length 13;

Best Local Similarity 88.9%; Pred. No. 1.5;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9

Db 3 VLHDDLLEA 11

RESULT 13

AD117193

ID AD117193 standard; protein; 1136 AA.

AC AD117193;

XX 15-APR-2004 (first entry)

DT

XX Human NOVX protein homologue SegID 729.

DE

XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;

KW inflammation; autoimmune disorder; allergy; blood disorder;

KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.

XX

OS Homo sapiens.

XX W0200268649-A2.

PN W0200268649-A2.

XX 06-SEP-2002.

PD

XX 31-JAN-2002; 2002WO-US002785.

PF

XX 31-JAN-2001; 2001US-0265395P.

PR

XX 31-JAN-2001; 2001US-0265412P.

PR

XX 31-JAN-2001; 2001US-0265514P.

PR

XX 02-FEB-2001; 2001US-0266406P.

PR

XX 05-FEB-2001; 2001US-0266767P.

PR

XX 07-FEB-2001; 2001US-0266975P.

PR

XX 08-FEB-2001; 2001US-0267057P.

PR

XX 09-FEB-2001; 2001US-0267823P.

PR

XX 15-FEB-2001; 2001US-0268974P.

PR

XX 26-FEB-2001; 2001US-0271664P.

PR

XX 27-FEB-2001; 2001US-0271839P.

PR

XX 02-MAR-2001; 2001US-0271855P.

PR

XX 14-MAR-2001; 2001US-0273046P.

PR

XX 14-MAR-2001; 2001US-0275925P.

PR

XX 14-MAR-2001; 2001US-0275950P.

PR

XX 14-MAR-2001; 2001US-0275989P.

PR

XX 15-MAR-2001; 2001US-0276448P.

PR

XX 15-MAR-2001; 2001US-0276397P.

PR

XX 16-MAR-2001; 2001US-0276768P.

PR

XX 20-MAR-2001; 2001US-0278652P.

PR

XX 26-MAR-2001; 2001US-0278775P.

PR

XX 29-MAR-2001; 2001US-0279882P.

PR

XX 29-MAR-2001; 2001US-0279884P.

PR

XX 30-MAR-2001; 2001US-0280147P.

PR

XX 11-APR-2001; 2001US-0282992P.

PR

XX 20-APR-2001; 2001US-02851083P.

PR

XX 23-APR-2001; 2001US-0285133P.

PR

XX 23-APR-2001; 2001US-0285749P.

PR

XX 03-MAY-2001; 2001US-0288327P.

PR

XX 29-MAY-2001; 2001US-0294047P.

PR

XX 30-MAY-2001; 2001US-0294473P.

PR

PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 13-JUN-2001; 2001US-0299324P.
PR 16-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0312908P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.

(CURA-) CURAGEN CORP.
PI Tchernev VT, Spytek KA, Zerhusen BD, Paturajan M, Shinkets RA;
PI Li L, Gangolli EA, Padigaru M, Anderson DM, Raattelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grose WM, Albrodook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706398/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 729; 1498bp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytosstatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antirheitic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virocidic, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
XX Sequence 1136 AA;
XX
XX Query Match 94.7%; Score 36; DB 5; Length 1136;
XX Best Local Similarity 88.9%; Pred. No. 1.6e+02;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 VLXDDLLEA 9
XX |||||
XX Db 137 VLRDDLEA 145
XX
XX RESULT 14
XX AD116744

ID AD116744 standard; protein; 1136 AA.
XX
XX AC AD116744;
XX
XX DT 15-APR-2004 (first entry)
XX
XX DE Human NOVX protein to treat human pathological conditions SeqID280.
XX
XX KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; AIDS; obesity; aschma;
KW acquired immunodeficiency syndrome; AIDS; obesity; aschma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
KW cytosstatic; cardiant; antiinflammatory; immunosuppressive; antiallergic;
KW haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;
KW antiasthmatic; nephrotropic; antirheitic; hepatotropic;
KW neuroprotective; nootropic; antibacterial; virocidic; antiparasitic;
KW relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; pharmacogenomic.
XX
XX OS Homo sapiens.
XX
XX PN WO200268649-A2.
XX
XX PD 06-SEP-2002.
XX
XX 31-JAN-2002; 2002WO-US002785.
XX
XX 31-JAN-2001; 2001US-0265395P.
XX 31-JAN-2001; 2001US-0265412P.
XX 31-JAN-2001; 2001US-0265514P.
XX 31-JAN-2001; 2001US-0265517P.
XX 02-FEB-2001; 2001US-0266406P.
XX 05-FEB-2001; 2001US-0266767P.
XX 07-FEB-2001; 2001US-0266975P.
XX 07-FEB-2001; 2001US-0267057P.
XX 08-FEB-2001; 2001US-0267459P.
XX 09-FEB-2001; 2001US-0267823P.
XX 15-FEB-2001; 2001US-0268974P.
XX 26-FEB-2001; 2001US-0271664P.
XX 27-FEB-2001; 2001US-0271839P.
XX 27-FEB-2001; 2001US-0271855P.
XX 02-MAR-2001; 2001US-0272788P.
XX 02-MAR-2001; 2001US-0273046P.
XX 14-MAR-2001; 2001US-0275925P.
XX 14-MAR-2001; 2001US-0275947P.
XX 14-MAR-2001; 2001US-0275950P.
XX 14-MAR-2001; 2001US-0275989P.
XX 15-MAR-2001; 2001US-0276448P.
XX 15-MAR-2001; 2001US-0276450P.
XX 16-MAR-2001; 2001US-0276397P.
XX 16-MAR-2001; 2001US-0276768P.
XX 20-MAR-2001; 2001US-0278652P.
XX 26-MAR-2001; 2001US-0278775P.
XX 26-MAR-2001; 2001US-0278778P.
XX 29-MAR-2001; 2001US-0279882P.
XX 29-MAR-2001; 2001US-0279884P.
XX 30-MAR-2001; 2001US-0280147P.
XX 11-APR-2001; 2001US-0282922P.
XX 11-APR-2001; 2001US-0283083P.
XX 20-APR-2001; 2001US-0285133P.
XX 23-APR-2001; 2001US-0285749P.
XX 03-MAY-2001; 2001US-0286327P.
XX 03-MAY-2001; 2001US-0288504P.
XX 29-MAY-2001; 2001US-0294047P.
XX 30-MAY-2001; 2001US-0294473P.
XX 08-JUN-2001; 2001US-0296959P.
XX 18-JUN-2001; 2001US-0296964P.
XX 19-JUN-2001; 2001US-0299324P.
XX 13-AUG-2001; 2001US-0312020P.
XX 16-AUG-2001; 2001US-0312889P.
XX 16-AUG-2001; 2001US-0312908P.
XX 21-AUG-2001; 2001US-0315470P.

PR 28-AUG-2001; 2001US-0315470P.
 PR 31-AUG-2001; 2001US-0316447P.
 PR 07-SEP-2001; 2001US-0318115P.
 PR 07-SEP-2001; 2001US-0318118P.
 PR 12-SEP-2001; 2001US-0318740P.
 PR 19-SEP-2001; 2001US-0323379P.
 PR 18-OCT-2001; 2001US-0330245P.
 PR 18-OCT-2001; 2001US-0330308P.
 PR 14-NOV-2001; 2001US-03332701P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Tchernev VT, Spytek KA, Zernhusen BD, Patumajan M, Shinkets RA;
 PI Li L, Ganggoli EA, Padigaru M, Anderson DM, Raetelli L, Miller CE;
 PI Gerlach VU, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CPA;
 PI Furtak K, Grose WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX
 DR MPI: 2002-706998/76.
 DR N-PSDB; ADI16743.
 XX
 PT New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; SEQ ID NO 280; 1496pp; English.
 XX
 CC This invention relates to a novel nucleic acid, and encoded polypeptides
 CC thereof, which have properties related to the stimulation of biochemical
 CC or physiological responses in a cell, tissue, organ or organism.
 CC Specifically, it refers to the use of biologically active fragments for
 CC diagnostic and prognostic assays and furthermore in the treatment of
 CC diverse pathological conditions. The present invention describes novel
 CC human and murine NOVX proteins, as well as methods to modulate their
 CC expression using antisense oligos, ribozymes and peptide nucleic acids.
 CC The NOVX polypeptides, polynucleotides and antibodies are useful in
 CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
 CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
 CC treating or preventing diseases such as inflammation, autoimmune
 CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
 CC (AIDS), obesity, asthma, immunoglobulin (Ig) A nephropathy, cirrhosis,
 CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
 CC and epilepsy. Accordingly, these molecules have many activities including
 CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
 CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
 CC antiaesthetic, nephroprotective, antiarthritic, hepatocrotic,
 CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
 CC relaxant and anticonvulsant. In addition, they are useful in screening
 CC assays to identify small molecules that modulate or inhibit, for example,
 CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
 CC used as in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. This polypeptide is a human NOVX protein of the
 CC invention.
 CC
 SQ Sequence 1136 AA;
 XX
 QY Query Match 94.7%; Score 36; DB 5; Length 1136;
 Best Local Similarity 88.9%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 1 VLXDDLEA 9
 137 VLRDDLEA 145
 XX
 XX
 XX
 AC ADH61299;
 XX
 DT 25-MAR-2004 (first entry)

DE INTSIG protein 7740249CD1, SEQ ID 16.
 XX
 XX Cytostatic; anorectic; immunosuppressive; gene therapy;
 KW INTSIG-agonist; INTSIG-agonist; human;
 KW intracellular signaling molecule; INTSIG; autoimmune disorder; obesity;
 KW cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO2004001005-A2.
 XX
 PD 31-DEC-2003.
 XX
 PF 18-JUN-2003; 2003WO-US019589.
 XX
 PR 19-JUN-2002; 2002US-0390470P.
 PR 16-AUG-2002; 2002US-0404235P.
 PR 12-SEP-2002; 2002US-0410424P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Li JK, Chawla NK, Gietzen KJ, Richardson TW, Kable AE;
 PI Ramkumar J, Elliott VS, Khare R, Ison CH, Bulloch SA, Yang J;
 PI Wilson AD, Griffin JA, Yang YG, Sprague WM, Tang YT;
 XX
 DR MPI: 2004-082488/08.
 DR N-PSDB; ADH61317.
 XX
 PT New human intracellular signaling molecule (INTSIG) polypeptide, useful
 PT for preparing a composition for treating a disease associated with
 PT decreased expression or overexpression of functional INTSIG e.g., cancer.
 XX
 PS Claim 1a; SEQ ID NO 16; 217pp; English.
 XX
 CC The invention relates to an isolated human intracellular signaling
 CC molecule (INTSIG) polypeptide. Also disclosed are the isolated
 CC polynucleotides encoding the polypeptides. The polypeptide is useful for
 CC preparing a composition for diagnosing or treating a disease or condition
 CC associated with decreased expression or overexpression of functional
 CC INTSIG e.g. autoimmune disorders, obesity or cancer. The current sequence
 CC represents an INTSIG polypeptide of the invention.
 CC
 SQ Sequence 1136 AA;
 XX
 QY Query Match 94.7%; Score 36; DB 8; Length 1136;
 Best Local Similarity 88.9%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 1 VLXDDLEA 9
 137 VLRDDLEA 145

Search completed: December 29, 2004, 23:09:37
 Job time : 155 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:06:35 ; Search time 38 Seconds
(without alignments)
15.707 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLRA 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	86.8	192	4	US-09-198-452A-404
2	32	84.2	635	4	US-09-252-991A-31013
3	32	84.2	4150	3	US-09-428-517-2
4	31	81.6	149	4	US-10-101-464A-529
5	30	78.9	454	4	US-09-489-039A-11396
6	30	78.9	776	4	US-09-252-991A-30380
7	30	78.9	822	4	US-09-886-319A-34
8	30	78.9	847	4	US-09-252-991A-29988
9	30	78.9	1410	3	US-09-335-409-3
10	30	78.9	1410	3	US-09-568-102-3
11	30	78.9	1410	3	US-09-567-969-3
12	30	78.9	1410	3	US-09-568-480-3
13	30	78.9	1410	3	US-09-568-486-3
14	30	78.9	1410	3	US-09-568-472-3
15	30	78.9	1410	3	US-09-567-899-3
16	29	76.3	142	4	US-09-248-796A-17805
17	29	76.3	199	4	US-09-543-681A-7115
18	29	76.3	259	4	US-09-252-991A-22086
19	29	76.3	263	4	US-09-252-991A-22624
20	29	76.3	410	4	US-09-489-039A-10283
21	29	76.3	413	4	US-09-489-039A-10299
22	29	76.3	415	4	US-09-583-110-4250
23	29	76.3	425	4	US-09-543-681A-7107
24	29	76.3	717	4	US-10-101-464A-810
25	29	76.3	1148	4	US-09-538-092-156
26	28	73.7	99	4	US-09-543-681A-8303
27	28	73.7	195	4	US-09-394-142B-16

28	28	73.7	251	5	PCT-US95-12357A-1	Sequence 1, Appli
29	28	73.7	252	4	US-09-506-286B-2	Sequence 2, Appli
30	28	73.7	252	4	US-09-506-286B-5	Sequence 5, Appli
31	28	73.7	252	4	US-09-762-861B-2	Sequence 2, Appli
32	28	73.7	252	4	US-09-762-861B-5	Sequence 5, Appli
33	28	73.7	252	4	US-10-065-133A-2	Sequence 2, Appli
34	28	73.7	252	4	US-10-065-133A-5	Sequence 5, Appli
35	28	73.7	293	4	US-09-934-899-2	Sequence 2, Appli
36	28	73.7	293	4	US-09-934-899-5	Sequence 5, Appli
37	28	73.7	296	4	US-09-583-110-5143	Sequence 22, Appli
38	28	73.7	334	4	US-09-489-039A-14107	Sequence 14107, A
39	28	73.7	550	1	US-08-674-168-29	Sequence 29, Appli
40	28	73.7	550	3	US-08-985-908-19	Sequence 19, Appli
41	28	73.7	550	3	US-08-985-908-4	Sequence 4, Appli
42	28	73.7	550	3	US-08-985-916-11	Sequence 11, Appli
43	28	73.7	605	4	US-09-270-767-46196	Sequence 46196, A
44	28	73.7	674	4	US-09-270-767-41618	Sequence 41618, A
45	28	73.7	679	4	US-09-270-767-44535	Sequence 44535, A

ALIGNMENTS

RESULT 1
US-09-198-452A-404
; Sequence 404, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 404
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-404

Query Match 86.8% ; Score 33; DB 4; Length 192;
Best Local Similarity 77.8% ; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 VLXDDLRA 9
DB 88 VLXDDLRA 96
RESULT 2
US-09-252-991A-31013
; Sequence 31013, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31013
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31013

Query Match 84.2%; Score 32; DB 4; Length 635;
Best Local Similarity 77.8%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
| | | | |
| | | | |
Db 119 VLADLLEA 127

RESULT 3

US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Bectelach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 84.2%; Score 32; DB 3; Length 4150;
Best Local Similarity 87.5%; Pred. No. 8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDILLEA 9
| | | | |
| | | | |
Db 1781 LRDILLEA 1788

RESULT 4

US-10-101-464A-529
; Sequence 529, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Eucalyptus grandis

US-10-101-464A-529
Query Match 81.6%; Score 31; DB 4; Length 149;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
| | | | |
| | | | |
Db 73 ILDELLEA 81

RESULT 5

US-09-489-039A-11396
; Sequence 11396, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11396
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11396

Query Match 78.9%; Score 30; DB 4; Length 454;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
| | | | |
| | | | |
Db 379 VLDDLLEA 387

RESULT 6

US-09-252-991A-30380
; Sequence 30380, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30380
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30380

Query Match 78.9%; Score 30; DB 4; Length 776;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDILLEA 9
| | | | |
| | | | |
Db 618 LADLLEA 625

RESULT 7

US-09-886-319A-34


```

; Sequence 34, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-319A-34

```

```

Query Match      78.9%; Score 30; DB 4; Length 822;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 VLXDDLE 8
      |||||
Db      552 VLKDDLE 559

```

```

RESULT 8
US-09-252-991A-29988
; Sequence 29988, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29988
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29988

```

```

Query Match      78.9%; Score 30; DB 4; Length 847;
Best Local Similarity 75.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 VLXDDLE 8
      |||||
Db      189 VLADDLVE 196

```

```

RESULT 9
US-09-335-409-3
; Sequence 3, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas

```

```

; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-335-409-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LXXDDLEA 9
      |||||
Db      474 LDDDDLEA 481

```

```

RESULT 10
US-09-568-102-3
; Sequence 3, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LXXDDLEA 9
      |||||
Db      474 LDDDDLEA 481

```

```

RESULT 11
US-09-567-969-3
; Sequence 3, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHLONES
; FILE REFERENCE: 4-30582A

```

```

; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-969-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LKDDLLEA 9
        |||:||||
Db      474 LKDDLLEA 481

```

```

RESULT 12
US-09-568-480-3
; Sequence 3, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LKDDLLEA 9
        |||:||||
Db      474 LKDDLLEA 481

```

```

RESULT 13
US-09-568-486-3
; Sequence 3, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LKDDLLEA 9
        |||:||||
Db      474 LKDDLLEA 481

```

```

RESULT 14
US-09-568-472-3
; Sequence 3, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-472-3

```

```

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 LKDDLLEA 9
        |||:||||
Db      474 LKDDLLEA 481

```

```

RESULT 15
US-09-567-899-3
; Sequence 3, Application US/09567899
; Patent No. 6383787
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,899
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

```

US-09-567-899-3

Query Match 78.9%; Score 30; DB 3; Length 1410;
 Best Local Similarity 75.0%; Pred. No. 6.3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKDDLLBA 9
 | | | | |
 DB 474 LKDDMLBA 481

Search completed: December 29, 2004, 23:14:16
 Job time : 39 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 23:09:46 ; Search time 146 Seconds
(without alignments)
22.175 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLEA 9

Scoring table: BLASTUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	94.7	9	US-10-623-176-1	Sequence 1, Appl1
2	36	94.7	9	US-10-623-176-2	Sequence 2, Appl1
3	36	94.7	9	US-10-623-176-10	Sequence 10, Appl1
4	36	94.7	9	US-10-791-217-1	Sequence 1, Appl1
5	36	94.7	9	US-10-791-217-2	Sequence 2, Appl1
6	36	94.7	9	US-10-791-217-5	Sequence 5, Appl1
7	36	94.7	10	US-10-623-176-24	Sequence 24, Appl1
8	36	94.7	10	US-10-623-176-43	Sequence 43, Appl1
9	36	94.7	10	US-10-623-176-44	Sequence 44, Appl1
10	36	94.7	12	US-10-623-176-65	Sequence 65, Appl1
11	36	94.7	13	US-10-623-176-74	Sequence 74, Appl1
12	36	94.7	13	US-10-623-176-76	Sequence 76, Appl1
13	36	94.7	17	US-10-791-217-14	Sequence 14, Appl1

14	36	94.7	13	17	US-10-791-217-16	Sequence 16, Appl1
15	36	94.7	17	15	US-10-623-176-33	Sequence 33, Appl1
16	36	94.7	17	15	US-10-623-176-57	Sequence 57, Appl1
17	36	94.7	17	15	US-10-623-176-69	Sequence 69, Appl1
18	36	94.7	19	15	US-10-623-176-70	Sequence 70, Appl1
19	36	94.7	21	15	US-10-623-176-71	Sequence 71, Appl1
20	36	94.7	22	15	US-10-623-176-53	Sequence 53, Appl1
21	36	94.7	22	15	US-10-623-176-63	Sequence 63, Appl1
22	36	94.7	23	15	US-10-623-176-72	Sequence 72, Appl1
23	36	94.7	23	15	US-10-623-176-55	Sequence 55, Appl1
24	36	94.7	25	15	US-10-623-176-67	Sequence 67, Appl1
25	36	94.7	26	15	US-10-623-176-66	Sequence 66, Appl1
26	36	94.7	27	15	US-10-623-176-51	Sequence 51, Appl1
27	36	94.7	27	15	US-10-623-176-60	Sequence 60, Appl1
28	36	94.7	28	15	US-10-623-176-61	Sequence 61, Appl1
29	36	94.7	29	15	US-10-623-176-49	Sequence 49, Appl1
30	36	94.7	29	15	US-10-623-176-59	Sequence 59, Appl1
31	36	94.7	1136	15	US-10-072-012-280	Sequence 280, App
32	36	94.7	1136	15	US-10-072-012-729	Sequence 729, App
33	36	94.7	1165	15	US-10-072-012-730	Sequence 730, App
34	35	92.1	408	16	US-10-437-963-114620	Sequence 61761, A
35	33	86.8	95	16	US-10-767-701-61761	Sequence 8607, Ap
36	33	86.8	185	14	US-10-156-761-8607	Sequence 404, App
37	33	86.8	192	15	US-10-289-762-404	Sequence 225678,
38	33	86.8	218	17	US-10-425-115-225678	Sequence 7046, A
39	33	86.8	297	15	US-10-425-114-73046	Sequence 225676,
40	33	86.8	317	17	US-10-425-115-225676	Sequence 99, Appl
41	33	86.8	499	15	US-10-312-273-99	Sequence 54798, A
42	33	86.8	499	15	US-10-282-122A-54798	Sequence 41, Appl
43	32	84.2	9	15	US-10-623-176-41	Sequence 42, Appl
44	32	84.2	9	15	US-10-623-176-42	Sequence 45, Appl
45	32	84.2	9	15	US-10-623-176-45	

ALIGNMENTS

RESULT 1
US-10-623-176-1
Sequence 1, Application US/10623176
Publicatoin No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.-J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT FILING DATE: 2003-07-18
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
OTHER INFORMATION: wherein X can be R or H
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(9)
US-10-623-176-1
Query Match 94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.Se+06;

```
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 2
US-10-623-176-2
; Sequence 2, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Elia A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-6047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-2

Query Match          94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(9)
US-10-623-176-10

Query Match          94.7%; Score 36; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 4
US-10-791-217-1
; Sequence 1, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Elia A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from histocompatibility antigen
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is HISTIDINE OR ARGININE RESIDUE
US-10-791-217-1

Query Match          94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
   |||||
Db 1 VLXDDLLEA 9

RESULT 5
US-10-791-217-2
; Sequence 2, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulimy, Elia A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
```

; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from histocompatibility antigen
US-10-791-217-2

Query Match 94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 6
US-10-791-217-5
; Sequence 5, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elie A.J.M.
; APPLICANT: Hunt, Donald F.
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-428505
; CURRENT APPLICATION NUMBER: US/10/791,217
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from KIAA0223 partial complementary DNA
US-10-791-217-5

Query Match 94.7%; Score 36; DB 17; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.5e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 7
US-10-623-176-24
; Sequence 24, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elie A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-604705
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165

; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(10)
US-10-623-176-24

Query Match 94.7%; Score 36; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 8
US-10-623-176-43
; Sequence 43, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elie A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-604705
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
; NAME/KEY: SITE
; LOCATION: (1)..(10)
US-10-623-176-43

Query Match 94.7%; Score 36; DB 15; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 1 VLHDDLLEA 9

RESULT 9
US-10-623-176-44
; Sequence 44, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elie A.J.M.
; APPLICANT: Hunt, Donald F.

```
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(10)
US-10-623-176-44
```

```
Query Match          94.7%; Score 36; DB 15; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 VLXDDLEA 9
        |||||||
Db      1 VLRDDLEA 9
```

```
RESULT 10
US-10-623-176-65
Sequence 65, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 65
LENGTH: 12
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: HA-1 peptide
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(12)
US-10-623-176-65
```

```
Query Match          94.7%; Score 36; DB 15; Length 12;
Best Local Similarity 88.9%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 VLXDDLEA 9
```

```
Db      2 VLRDDLEA 10
        |||||||
```

```
RESULT 11
US-10-623-176-74
Sequence 74, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 74
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: KIAA0223
OTHER INFORMATION: sequence derived from a presumed HA-1 negative
FEATURE:
NAME/KEY: SITE
LOCATION: (1)..(13)
US-10-623-176-74
```

```
Query Match          94.7%; Score 36; DB 15; Length 13;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 VLXDDLEA 9
        |||||||
Db      3 VLRDDLEA 11
```

```
RESULT 12
US-10-623-176-76
Sequence 76, Application US/10623176
Publication No. US20040092446A1
GENERAL INFORMATION:
APPLICANT: Goulmy, Els A.J.M.
APPLICANT: Hunt, Donald F.
APPLICANT: Engelhard, Victor H.
TITLE OF INVENTION: HA-1 epitopes and uses thereof
FILE REFERENCE: 2183-6047US
CURRENT APPLICATION NUMBER: US/10/623,176
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: 09/489,760
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: EP 97202303.0
PRIOR FILING DATE: 1997-07-23
PRIOR APPLICATION NUMBER: PCT/NL98/00424
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: JP 2000-504165
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 76
LENGTH: 13
TYPE: PRT
```



```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: KIAA0223
; OTHER INFORMATION: sequence derived from a presumed HA-1 Homozygous
; OTHER INFORMATION: positive individual
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(13)
US-10-623-176-76

```

```

Query Match          94.7%; Score 36; DB 15; Length 13;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 VLXDDLLEA 9
    |||||
Db 3 VLHDDLLEA 11

```

```

RESULT 13
US-10-791-217-14
; Sequence 14, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elia A.J.M
; APPLICANT: Hunt, Donald F
; APPLICANT: Engelhard, Victor H
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; PRIOR FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from HA-1/- phenotype
US-10-791-217-14

```

```

Query Match          94.7%; Score 36; DB 17; Length 13;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 VLXDDLLEA 9
    |||||
Db 3 VLHDDLLEA 11

```

```

RESULT 14
US-10-791-217-16
; Sequence 16, Application US/10791217
; Publication No. US20040191268A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elia A.J.M
; APPLICANT: Hunt, Donald F
; APPLICANT: Engelhard, Victor H
; TITLE OF INVENTION: The HA-1 Antigen
; FILE REFERENCE: 2183-4285US
; CURRENT APPLICATION NUMBER: US/10/791,217
; PRIOR FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16

```

```

; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: obtained from HA-1/+ phenotype
US-10-791-217-16

```

```

Query Match          94.7%; Score 36; DB 17; Length 13;
Best Local Similarity 88.9%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 VLXDDLLEA 9
    |||||
Db 3 VLHDDLLEA 11

```

```

RESULT 15
US-10-623-176-33
; Sequence 33, Application US/10623176
; Publication No. US20040092446A1
; GENERAL INFORMATION:
; APPLICANT: Goulmy, Elia A.J.M.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Engelhard, Victor H.
; TITLE OF INVENTION: HA-1 epitopes and uses thereof
; FILE REFERENCE: 2183-5047US
; CURRENT APPLICATION NUMBER: US/10/623,176
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/489,760
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: EP 97202303.0
; PRIOR FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: PCT/NL98/00424
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: JP 2000-504165
; PRIOR FILING DATE: 2000-01-24
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HA-1
; OTHER INFORMATION: polymorphic region wherein X can be H or R
; NAME/KEY: SITE
; LOCATION: (1)..(17)
US-10-623-176-33

```

```

Query Match

```

```

Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 VLXDDLLEA 9
    |||||
Db 7 VLHDDLLEA 15

```

```

Search completed: December 29, 2004, 23:24:41
Job time : 147 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:58:45 ; Search time 40 Seconds
(without alignments)
21.649 Million cell updates/sec

Title: US-10-791-217A-1

Perfect score: 38

Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	1165	2	DS9433
2	34	89.5	616	2	C69226
3	33	86.8	174	2	C89075
4	33	86.8	499	2	G85038
5	33	86.8	499	2	G72083
6	32	84.2	112	2	S31180
7	32	84.2	797	2	T50072
8	31	81.6	276	2	AD1291
9	31	81.6	276	2	AB1663
10	31	81.6	289	2	C81175
11	31	81.6	289	2	DB1930
12	31	81.6	397	2	A28180
13	31	81.6	497	2	G87793
14	31	81.6	750	2	B83794
15	31	81.6	1042	2	T13647
16	30	78.9	302	2	T26975
17	30	78.9	302	2	AB3246
18	30	78.9	361	2	T48014
19	30	78.9	623	2	E75221
20	30	78.9	683	2	DB2674
21	30	78.9	822	2	I38728
22	30	78.9	962	2	JT0669
23	30	78.9	1699	2	AB2018
24	30	78.9	1999	2	T14165
25	29	76.3	101	2	E97582
26	29	76.3	101	2	AE2803
27	29	76.3	119	2	T35859
28	29	76.3	137	1	Q3BCEA
29	29	76.3	137	1	Q3BCEA

30	29	76.3	137	2	A10575	conserved hypothet
31	29	76.3	137	2	D90708	hypothetical prote
32	29	76.3	137	2	H85558	hypothetical prote
33	29	76.3	154	1	R5H530	ribosomal protein
34	29	76.3	156	2	T20267	hypothetical prote
35	29	76.3	178	2	D72293	hypothetical prote
36	29	76.3	192	2	A12271	transcription regu
37	29	76.3	207	2	G83538	probable oxidoredu
38	29	76.3	231	2	B65140	hypothetical 26.3
39	29	76.3	231	2	H91164	hypothetical prote
40	29	76.3	231	2	H86010	hypothetical prote
41	29	76.3	244	2	D86725	hypothetical prote
42	29	76.3	251	2	S45438	cytochrome-c oxida
43	29	76.3	277	2	D69158	sensory transducti
44	29	76.3	292	2	S72323	ubiquinol-cytochro
45	29	76.3	322	2	S44847	K06f7.7 protein -

ALIGNMENTS

RESULT 1
DS9433
C. elegans protein Z37093 homolog [imported] - human
C.Species: Homo sapiens (man)
C.Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C.Accession: DS9433; E59433
R.Nagase, T.; Seki, N.; Ishikawa, K.; Ohira, M.; Kawarabayashi, Y.; Ohara, O.; Tanaka, A.
DNA Res. 3, 321-329, 1996
A.Title: Prediction of the coding sequences of unidentified human genes. VI. The coding
A.Reference number: DS9433; MUID:97191544; PMID:9039502
A.Accession: DS9433
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-1165 <NMG>
A.Cross-references: UNIPROT:Q92619; GB:BA13212; PID:g1504026; PIDN:BA13212.1
R.Ohara, O.; Nagase, T.; Kikuno, R.; Nomura, N.
Submitted to Genbank, August 1996
A.Reference number: E59433
A.Accession: E59433
A.Status: Preliminary
A.Molecule type: DNA
A.Residues: 1-1165 <STO2>
A.Cross-references: GB:BA13212; PID:g1504026; PIDN:BA13212.1
Query Match
Best Local Similarity 94.7%; Score 36; DB 2; Length 1165;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLEA 9
DB 166 VLXDDLEA 174
RESULT 2
C69226
Type I restriction modification enzyme, subunit M - Methanobacterium thermoautotrophicum
C.Species: Methanobacterium thermoautotrophicum
C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C.Accession: C69226
R.Smith, D.R.; Doucette-Stamm, L.A.; Delongchery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A.Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A.Reference number: A59000; MUID:98037514; PMID:9371463
A.Accession: C69226
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-616 <MT1>
A.Cross-references: UNIPROT:O27025; GB:AE000868; GB:AE000666; NID:g2622025; PIDN:AA8544
A.Experimental source: strain Delta H

C:Genetics:
A:Gene: MTH942
A:Start codon: GTG
C:Superfamily: type I site-specific deoxyribonuclease chain hcdm

Query Match 89.5%; Score 34; DB 2; Length 616;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||:||||
DB 465 VLXDDLLEA 473

RESULT 3
C89075
protein K04A8.9 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89075
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A15000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Accession: C89075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: UNIPROT:Q94256; GB:chr_V; PIDN:AAc48055.1; PID:g1658358; GSPDB:GN000
C:Genetics:
A:Gene: K04A8.9
A:Map position: 5

Query Match 86.8%; Score 33; DB 2; Length 174;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||:||||
DB 129 VLXDDLLEA 137

RESULT 4
G86538
leucyl aminopeptidase A [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G86538
R:Shitai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishl, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: G86538
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-499 <STO>
A:Cross-references: UNIPROT:Q928F8; GB:BA000008; NID:g8978757; PIDN:BAa98593.1; GSPDB:GN
C:Genetics:
A:Gene: pepA
C:Superfamily: Cytosol aminopeptidase

Query Match 86.8%; Score 33; DB 2; Length 499;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||:||||
DB 395 VLXDDLLEA 403

RESULT 5
G72083
leucine aminopeptidase CP0370 [imported] - Chlamydia pneumoniae (strains CML029 and

N:Alternate names: leucyl aminopeptidase A
C:Species: Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: G72083; F81583
R:Kauman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: G72083
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-499 <ARN>
A:Cross-references: UNIPROT:Q928F8; GB:AE001623; GB:AE001363; NID:g4376662; PIDN:AD1852
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Grimm, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOpn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81583
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-499 <RNA>
A:Cross-references: GB:AE002199; GB:AE002161; NID:g7189293; PIDN:AAf38219.1; PID:g718929
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: pepA; CP0370
C:Superfamily: Cytosol aminopeptidase

Query Match 86.8%; Score 33; DB 2; Length 499;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
|||:||||
DB 395 VLXDDLLEA 403

RESULT 6
S33180
nitrogen regulatory protein P-II - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S33180
R:Zinchenko, V.V.; Churin, U.N.; Shestopalov, V.I.; Shestakov, S.V.
submitted to the EMBL Data Library, April 1993
A:Description: Nucleotide sequence and characterization of glnBA operons of the purple p
A:Reference number: S33180
A:Accession: S33180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <ZIN>
A:Cross-references: UNIPROT:P43519; EMBL:X71659; NID:g297112; PIDN:CAa50650.1; PID:g8097
A:Experimental source: strain 2R
C:Genetics:
A:Gene: glnB
A:Start codon: GTG
C:Function:
A:Description: important for the control of glutamine synthase; in nitrogen-limiting con
to form P-II-UWP; P-II-UWP allows the deadenylation of glutamine synthase, thus activa
GS
A:Note: P-II indirectly controls the transcription of the GS gene (glnA) by preventing N

Query Match 84.2%; Score 32; DB 2; Length 112;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
|||:|
Db 64 VLADDMVER 72

RESULT 7
T50072
hypothetical protein SPAC1486.03c [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50072
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.
submitted to the EMBL Data Library, December 1999
A:Reference number: 225034
A:Accession: T50072
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-797 <MCD>
A:Cross-references: UNIPROT:Q9UTK6; EMBL:AL133357; PIDN:CAB62413.1; GSPDB:GN00066; SPDB:
C:Genetics:
A:Gene: SPDB:SPAC1486.03c
A:Map position: 1

Query Match 84.2%; Score 32; DB 2; Length 797;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
|||:|
Db 81 VTADDLLEA 89

RESULT 8
AD1291
sugar ABC transporter, permease protein homolog lmo1732 [imported] - *Listeria monocytogenes*
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1291
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fahn, H.
Science 294, 849-852, 2001
D.; Jones, L.M.; Karsl, U.
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.; Ma
ok, C.; Schueter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1291
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <GLA>
A:Cross-references: UNIPROT:Q8Y6F5; GB:NC_003210; PIDN:CAC99810.1; PID:G16411186; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1732
C:Superfamily: maltose transport protein malG

Query Match 81.6%; Score 31; DB 2; Length 276;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKDDLLEA 9
:|||||
Db 165 ISDDLLEA 172

RESULT 9
AB1663
sugar ABC transporter, permease protein homolog lln1843 [imported] - *Listeria innocua* (S
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1663

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fahn, H.
Science 294, 849-852, 2001
D.; Jones, L.M.; Karsl, U.
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.; Ma
ok, C.; Schueter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1663
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <GLA>
A:Cross-references: UNIPROT:Q92AS6; GB:AL592022; PIDN:CAC97074.1; PID:G16414345; GSPDB:(
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lln1843
C:Superfamily: maltose transport protein malG

Query Match 81.6%; Score 31; DB 2; Length 276;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LKDDLLEA 9
:|||||
Db 165 ISDDLLEA 172

RESULT 10
CB1175
UTP-glucose-1-phosphate uridylyltransferase NMB0638 [imported] - *Neisseria meningitidis*
C:Species: *Neisseria meningitidis*
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: CB1175
R:Retcelijn, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Bisen, J.P.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.,
.; H.; Qin, H.; Vamathavan, U.; Gill, J.; Scarlato, V.; Weisigant, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vc
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: CB1175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <TEP>
A:Cross-references: UNIPROT:Q9K0G7; GB:AE002419; GB:AE002098; NID:G7225863; PIDN:AAF4106
C:Genetics:
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0638
C:Superfamily: *Escherichia coli* UTP-glucose-1-phosphate uridylyltransferase

Query Match 81.6%; Score 31; DB 2; Length 289;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
|||:|
Db 130 TLADDLLEA 138

RESULT 11
DB1930
probable UTP-glucose-1-phosphate uridylyltransferase NMA0848 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: DB1930
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
.; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: DB1930
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-289 <PAR>
 A:Cross-references: UNIPROT:Q9JVC6; GB:AL162754; GB:AL157959; NID:G7379424; PIDN:CB8412
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: galY; NMA0848
 C:Superfamily: Escherichia coli UTP-glucose-1-phosphate uridylyltransferase

Query Match
 Best Local Similarity 81.6%; Score 31; DB 2; Length 289;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 : |||||
 Db 130 ILADLLEA 138

RESULT 12
 A28180
 carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) large chain - Methanos
 N:Alternate names: carbamoyl-phosphate synthetase (glutamine-hydrolyzing), ammonia chain
 C:Species: Methanococcus barkeri
 C>Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 09-Jul-2004
 C:Accession: A28180
 R:Morris, C.J.; Reeve, J.N.
 J. Bacteriol. 170, 3125-3130, 1988
 A:Title: Conservation of structure in the human gene encoding argininosuccinate synthetase
 A:Reference number: A91873; MUID:88257029; PMID:3133361
 A:Accession: A28180
 A:Molecule type: DNA
 A:Residues: 1-397 <MOR>
 A:Cross-references: UNIPROT:P13258
 C:Genetics:
 A:Gene: canB
 C:Keywords: arginine biosynthesis; ligase; pyrimidine nucleotide biosynthesis
 F:1-308/Domain: diocin carboxylase homology <ECH>

Query Match
 Best Local Similarity 81.6%; Score 31; DB 2; Length 397;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 : |||||
 Db 67 ILIDPLEA 75

RESULT 13
 G87793
 protein C27A12.7 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C:Accession: G87793
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A>Note: see website genome.wustl.edu/gsc/C.elegans/ and www_sanger.ac.uk/Projects/C_elg
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: G87793
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-497 <STO>
 A:Cross-references: UNIPROT:O01964; GB:chr_I; PIDN:AAB93644.1; PID:G2105479; GSPDB:GN000
 C:Genetics:
 A:Gene: C27A12.7
 A:Map position: 1

Query Match
 Best Local Similarity 81.6%; Score 31; DB 2; Length 497;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 : |||||
 Db 55 VLIDNLEA 63

RESULT 14
 B83794
 hypothetical protein BH1154 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: B83794
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeki, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: B83794
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-750 <STO>
 A:Cross-references: UNIPROT:Q9K007; GB:AP001511; GB:BA000004; NID:G10173727; PIDN:BA8048
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1154

Query Match
 Best Local Similarity 81.6%; Score 31; DB 2; Length 750;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 : |||||
 Db 581 ILIDPLEA 589

RESULT 15
 T13647
 hypothetical protein 95B7.8 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
 C:Accession: T13647
 R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
 submitted to the EMBL Data Library, April 1999
 A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
 A:Reference number: Z17694
 A:Accession: T13647
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1042 <FER>
 A:Cross-references: UNIPROT:O76911; EMBL:AL021728; PIDN:CAA16820.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0000376
 A:Introns: 94/3; 833/3; 937/3
 A>Note: EG:95B7.8
 C:Superfamily: fruit fly hypothetical protein EG_95B7.8

Query Match
 Best Local Similarity 81.6%; Score 31; DB 2; Length 1042;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 : |||||
 Db 595 VLFDLLEA 603

Search completed: December 29, 2004, 23:13:34
 Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 22:57:40 ; Search time 188 Seconds.

(without alignments)
27.545 Million cell updates/sec

Title: US-10-791-217A-1
Perfect score: 38
Sequence: 1 VLXDDLLEA 9

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	94.7	668	06LZ81	Q6LZ81 methanococ
2	36	94.7	668	CAF30304	Q6LZ81 methanoco
3	36	94.7	1131	081YN3	Q81YN3 homo sapien
4	36	94.7	1136	06P189	Q6P189 homo sapien
5	36	94.7	1136	08HX84	Q8HX84 homo sapien
6	36	94.7	1136	AAH65223	AAH65223 homo sapi
7	35	92.1	1165	Q92619	Q92619 homo sapien
8	35	92.1	585	08EJ89	Q8EJ89 shewanella
9	34	89.5	146	074B30	Q74B30 geobacter s
10	34	89.5	146	AAH35828	AAH35828 geobacter
11	34	89.5	376	Q8WV81	Q8WV81 ixodes scap
12	34	89.5	616	Q27025	Q27025 methanobact
13	33	86.8	174	Q94256	Q94256 caenorhabdi
14	33	86.8	185	Q82P65	Q82P65 streptomyce
15	33	86.8	291	081SV5	Q81SV5 bacillus an
16	33	86.8	291	AAH35306	AAH35306 bacillus
17	33	86.8	292	Q6HL66	Q6HL66 bacillus th
18	33	86.8	292	Q61163	Q61163 bacillus an
19	33	86.8	292	081FT0	Q81FT0 bacillus ce
20	33	86.8	499	1 AMPA_CHLPN'	Q928F8 chlamydia p
21	32	84.2	112	1 GMB_RHOSH	P43519 rhodobacter
22	32	84.2	112	031158	Q31158 rhodobacter
23	32	84.2	436	Q8EP44	Q8EP44 oceanobacti
24	32	84.2	616	Q50407	Q50407 mycobacteri
25	32	84.2	797	Q9UTK6	Q9UTK6 schizosacch
26	32	84.2	800	Q8FTX1	Q8FTX1 corynobacte
27	32	84.2	894	2 Q9FMC7	Q9FMC7 oryza sativ
28	32	84.2	920	2 Q7RYV8	Q7RYV8 neurospora
29	32	84.2	1155	2 Q85ZFS	Q85ZFS oryza sativ
30	32	84.2	2041	2 Q7S374	Q7S374 neurospora
31	32	84.2	4150	2 Q9KIV4	Q9KIV4 streptomyce

32	31	81.6	169	2	Q73B06	Q73B06 bacillus ce
33	31	81.6	169	2	AA40543	AA40543 bacillus
34	31	81.6	226	2	Q7OMX4	Q7OMX4 anophelis g
35	31	81.6	263	2	Q98L75	Q98L75 rhizobium l
36	31	81.6	276	2	Q92A86	Q92A86 listeria in
37	31	81.6	276	2	Q816F5	Q816F5 listeria mo
38	31	81.6	276	2	Q71YT6	Q71YT6 listeria mo
39	31	81.6	276	2	AA704528	AA704528 listeria
40	31	81.6	289	2	Q9JVG6	Q9JVG6 neisseria m
41	31	81.6	289	2	Q9K0G7	Q9K0G7 neisseria m
42	31	81.6	293	2	Q8H305	Q8H305 oryza sativ
43	31	81.6	299	2	Q83VQ1	Q83VQ1 methylolact
44	31	81.6	306	1	ISH1_RHOPA	Q6N3G0 rhodospheudo
45	31	81.6	306	2	CAE29175	CAE29175 rhodospheu

ALIGNMENTS

RESULT 1						
06LZ81	PRELIMINARY;	PRT;	668	AA.		
AC	Q6LZ81;					
DT	05-JUL-2004 (TREMBlrel. 27, Created)					
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)					
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)					
DE	MCM family related protein.					
GN	OrderedLocustNames=MMP0748;					
OS	Methanococcus maripaludis.					
OC	Archaea; Euryarchaeota; Methanococci; Methanococcales;					
OC	Methanococcaceae; Methanococcus.					
OX	NCBI_TaxID=39152;					
RP	[1]					
RE	SEQUENCE FROM N.A.					
RC	STRAIN=S2 / LI;					
RA	Hendrickson E.L., Kaul R., Zhou Y., Boyse D., Chapman P., Chung J.,					
RA	Conway de Macario E., Dodsworth J., Gillett W., Graham D.E.,					
RA	Haydock A.K., Kang A., Land M.L., Levy R., Lie T.J., Major T.,					
RA	Moore B., Porat I., Overbeek R., Palmeiri A., Rouse G.,					
RA	Saenphimmachak C., Soll D., Whitman W.B., Larimer F.W., Olson M.V.,					
RA	Leigh J.A.;					
RT	"Complete genome sequence of the mesophilic hydrogenotrophic					
RT	methanogen Methanococcus maripaludis."					
RL	Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.					
DR	EMBL; BX957221; CAF30304.1; -					
DR	InterPro; IPR002048; EF-hand.					
DR	InterPro; IPR001208; MCM.					
DR	Pfam; PF00493; MCM; 1.					
DR	PRINTS; PR01657; MCMFAMILY.					
DR	ProDom; PD001041; MCM; 1.					
DR	SMART; SM00350; MCM; 1.					
DR	PROSITE; PS00018; EF_HAND; UNKNOWN_1.					
DR	PROSITE; PS50051; MCM_2; 1.					
KW	Complete proteome.					
SC	SEQUENCE 668 AA; 75660 MW; 2ADAF380B1049F4 CRC64;					
Query Match						
Best Local Similarity			94.7%;	Score 36;	DB 2;	Length 668;
Matches			8;	Conservative	0;	Mismatches 1;
						Indels 0;
						Gaps 0;
Qy	1	VLXDDLLEA 9				
Db	628	VLXDDLLEA 636				
RESULT 2						
CAF30304	PRELIMINARY;	PRT;	668	AA.		
ID	CAF30304					
AC	CAF30304;					
DT	03-MAR-2004 (TREMBlrel. 27, Created)					
DT	03-MAR-2004 (TREMBlrel. 27, Last sequence update)					
DT	04-MAY-2004 (TREMBlrel. 27, Last annotation update)					
DE	MCM family related protein.					

GN MMP0748.
 OS Methanococcus maripaludis.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanococcaceae; Methanococcus.
 NC NCB1_TaxId=39152;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=52 / L1;
 RA Hendrickson E.L., Kaul R., Zhou Y., Boyee D., Chapman P., Chung J.,
 RA Conway de Macario E., Dodsworth J., Gilllett W., Graham D.E.,
 RA Haydock A.K., Kang A., Land M.L., Levy R., Lie T.J., Major T.,
 RA Moore B., Port I., Overbeek R., Palmeiri A., Rouse G.,
 RA Saepthummachak C., Seell D., Whitman W.B., Larimer F.W., Olson M.V.,
 RA Leigh J.A.;
 RT "Complete genome sequence of the mesophilic hydrogenotrophic
 RT Methanogen Methanococcus maripaludis";
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, EX957221; CAF30304.1; --
 SQ SEQUENCE 668 AA; 75660 MW; 2ADAF380B1049F4 CRC64;

Query Match 94.7%; Score 36; DB 2; Length 668;
 Best Local Similarity 88.9%; Pred. No. 1;le+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 DB 628 VLXDDLLEA 636

RESULT 3
 ID Q8IYN3 PRELIMINARY; PRT; 1131 AA.
 AC Q8IYN3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Similar to PTP1-associated RhogAP 1 (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC035564; AAH35564.1; --
 DR HSRF, Q88935; 1F7C.
 DR GO; GO:0007242; P:intracellular signaling cascade, IEA.
 DR InterPro; IPR001060; Cdc15_Fes_CIP4.
 DR InterPro; IPR000345; Cytc_Feme_BS.
 DR InterPro; IPR002219; DAG_Fe-bind.
 DR InterPro; IPR000198; RhogAP.
 DR InterPro; IPR008936; RhogAP.
 DR Pfam; PF00130; C1_1; 1.
 DR Pfam; PF00620; RhogAP; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00055; FCH; 1.
 DR SMART; SM00324; RhogAP; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_Fe_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE; PS50081; DAG_Fe_BIND_DOM_2; 1.
 DR PROSITE; PS50238; RhogAP; 1.
 FT NON TER 1
 SQ SEQUENCE 1131 AA; 124081 MW; A606B1AA52014925 CRC64;

Query Match 94.7%; Score 36; DB 2; Length 1131;
 Best Local Similarity 88.9%; Pred. No. 2e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 DB 132 VLXDDLLEA 140

RESULT 4
 ID Q6P189 PRELIMINARY; PRT; 1136 AA.
 AC Q6P189;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Minor histocompatibility antigen HA-1.
 GN Name=HA-1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC065223; AAH65223.1; --
 DR InterPro; IPR001060; Cdc15_Fes_CIP4.
 DR InterPro; IPR000345; Cytc_Feme_BS.
 DR InterPro; IPR002219; DAG_Fe-bind.
 DR InterPro; IPR000198; RhogAP.
 DR InterPro; IPR008936; RhogAP.
 DR Pfam; PF00130; C1_1; 1.
 DR Pfam; PF00620; RhogAP; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00055; FCH; 1.
 DR SMART; SM00324; RhogAP; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_Fe_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE; PS50081; DAG_Fe_BIND_DOM_2; 1.
 DR PROSITE; PS50238; RhogAP; 1.
 SQ SEQUENCE 1136 AA; 124550 MW; 14B7073296102DD5 CRC64;

Query Match 94.7%; Score 36; DB 2; Length 1136;
 Best Local Similarity 88.9%; Pred. No. 2e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 DB 137 VLXDDLLEA 145

RESULT 5
 ID Q8HX84 PRELIMINARY; PRT; 1136 AA.

AC Q8HX84;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Minor histocompatibility antigen HA-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20549028; PubMed=11095984;
 RA MEDLINE W.E., Piehler A., Schmitz G.;
 RT "Genomic organization of the human cholesterol-responsive ABC
 RT transporter ABCA7: tandem linkage with the minor histocompatibility
 RT antigen HA-1 gene."
 RL Biochem. Biophys. Res. Commun. 278: 782-789 (2000).
 DR EMBL, AF308045; AAN04658.1; JOINED.
 DR EMBL, AF308046; AAN04658.1; JOINED.
 DR EMBL, AF308047; AAN04658.1; JOINED.
 DR EMBL, AF308048; AAN04658.1; JOINED.
 DR EMBL, AF308049; AAN04658.1; JOINED.
 DR EMBL, AF308050; AAN04658.1; JOINED.
 DR EMBL, AF308051; AAN04658.1; JOINED.
 DR EMBL, AF308052; AAN04658.1; JOINED.
 DR EMBL, AF308053; AAN04658.1; JOINED.
 DR EMBL, AF308054; AAN04658.1; JOINED.
 DR EMBL, AF308055; AAN04658.1; JOINED.
 DR EMBL, AF308056; AAN04658.1; JOINED.
 DR EMBL, AF308057; AAN04658.1; JOINED.
 DR EMBL, AF308058; AAN04658.1; JOINED.
 DR EMBL, AF308059; AAN04658.1; JOINED.
 DR EMBL, AF308060; AAN04658.1; JOINED.
 DR EMBL, AF308061; AAN04658.1; JOINED.
 DR EMBL, AF308062; AAN04658.1; JOINED.
 DR EMBL, AF308063; AAN04658.1; JOINED.
 DR EMBL, AF308064; AAN04658.1; JOINED.
 DR EMBL, AF308065; AAN04658.1; JOINED.
 DR EMBL, AF311102; AAN04658.1; JOINED.
 DR HSSP; Q98935; 1F7C.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR001060; Cdc15_Fes_CIP4.
 DR InterPro; IPR000345; CycC_Heme_BS.
 DR InterPro; IPR002219; DAG_FR-bind.
 DR InterPro; IPR000198; RhoGAP.
 DR InterPro; IPR008936; Rho_GAP.
 DR Pfam; PF00130; C1_1; 1.
 DR Pfam; PF00620; RhoGAP; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00055; FCH; 1.
 DR SMART; SM00324; RhoGAP; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_FR_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE; PS50081; DAG_FR_BIND_DOM_2; 1.
 DR PROSITE; PS50238; RHO_GAP_1.
 SQ SEQUENCE 1136 AA; 124613 MW; 9536787B3B1E1ED CRC64;

DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
 DE Minor histocompatibility antigen HA-1.
 GN HA-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalski U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RT Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC065223; AAH65223.1;
 SQ SEQUENCE 1136 AA; 124550 MW; 14B7073296102DD5 CRC64;

Query Match 94.7%; Score 36; DB 2; Length 1136;
 Best Local Similarity 88.9%; Pred. No. 2e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
 |||||
 Db 137 VLXDDLLEA 145

RESULT 7
 ID 092619 PRELIMINARY; PRT; 1165 AA.
 AC 092619;
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE KIAA0223 protein (Fragment).
 GN Name=KIAA0223;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=bone marrow;
 RX MEDLINE=97191544; PubMed=9039502;
 RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
 RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. VI.
 RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
 RT analysis of cDNA clones from cell line KG-1 and brain."
 RL DNA Res. 3:321-329 (1996).
 DR EMBL; D86976; BAA13212.1;
 DR PIR; D59433; D59433.

```

DR GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001060; Cdc15_Fes_CIP4.
DR InterPro; IPR000345; CycC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF00130; C1_1; 1.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00055; FCH; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50238; RHO GAP; 1.
DR NON_TER 1
SQ SEQUENCE 1165 AA; 127344 MW; 92EF768CAF58C9 CRC64;

Query Match
Best Local Similarity 94.7%; Score 36; DB 2; Length 1165;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
Db 166 VLXDDLLEA 174

RESULT 8
OBS99 PRELIMINARY; PRT; 585 AA.
ID OBS99;
AC OBS99;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 23, Last annotation update)
DE Type I restriction-modification system, M subunit.
GN Name-hsdM-1; OrderedLocustNames=SO0383;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beaman M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA Lee K., Berry K.J., Lee C., Muehlner J., Khouli H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealeon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:118-123 (2002).
DR EMBL; AE015486; AAN53466.1; -.
DR TIGR; S00383; -.
DR GO:0003677; P:DNA binding; IEA.
DR GO:0008170; F:N-methyltransferase activity; IEA.
DR GO:0008007; F:site-specific DNA-methyltransferase (adenin. .); IEA.
DR GO:0006306; P:DNA methylation; IEA.
DR GO:0006304; P:DNA modification; IEA.
DR InterPro; IPR004546; HsdM.
DR InterPro; IPR003665; Methylase_M.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR003356; N6_DNA_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF02506; Methylase_M; 1.
DR Pfam; PF02384; N6_Mtase; 1.
DR PRINTS; PR00507; N12N6MTPRASE.
DR TIGRFAW; TIGR00497; hsdM; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.

```

```

KW Complete proteome.
SQ SEQUENCE 585 AA; 66321 MW; 349D45BF52C98D0C CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 2; Length 585;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
Db 335 ILMDLLEA 343

RESULT 9
OBS30 PRELIMINARY; PRT; 146 AA.
ID OBS30;
AC OBS30;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=GSU2455;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beaman M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Davidson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouli H.M., Feldblyum T.V., Uterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments."
RL Science 302:1967-1969 (2003).
DR EMBL; AE017180; AAR35828.1; -.
DR TIGR; GSU2455; -.
DR InterPro; IPR005361; UPF0158.
DR Pfam; PF0191405; UPF0158; 1.
DR Hypothetical protein.
SQ SEQUENCE 146 AA; 17581 MW; 30B46BEE69DE5255 CRC64;

Query Match
Best Local Similarity 89.5%; Score 34; DB 2; Length 146;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLXDDLLEA 9
Db 9 ILMDLLEA 17

RESULT 10
AAR35828 PRELIMINARY; PRT; 146 AA.
ID AAR35828;
AC AAR35828;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN GSU2455.
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,

```

RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beaman M.J., Dodson R.J.,
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
 RA Gwinn M.L., Kolony J.F., Sullivan S.A., Haft D.H., Selengut J.,
 RA Davidson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
 RA Weidman J.E., Khoult H.M., Feldblyum T.V., Uterback T.R.,
 RA Aken S.E., Lovley D.R., Fraser C.M.,
 RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
 environments";
 RL Science 302:1967-1969 (2003).
 DR EMBL: AE017215; AAR35828.1; --
 KW TIGR; GSU2455; --
 SO SEQUENCE 146 AA; 17581 MW; 30846BEE69DE5255 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 146;
 Best Local Similarity 77.8%; Pred. No. 60;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLXDDLLEA 9
 Db 9 ILMDLLEA 17

RESULT 11
 Q8WB1 PRELIMINARY; PRT; 376 AA.
 ID Q8WB1;
 AC Q8WB1;
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DS Putative serine protease with signal anchor.
 OS Ixodes scapularis (Black-legged tick) (Deer tick).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
 OX NCBI_TaxId=6945;
 RN [1]
 RS SEQUENCE FROM N.A.
 RC STRAIN=Rhode Island; TISSUE=Salivary gland;
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
 RA Mather T.N., Ribeiro J.M.C.;
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Belongs to peptidase family 51.
 DR EMBL: AF483729; AAM93651.1; --
 DR GO: GO:0004263; F:cytochrome-c oxidase activity; IEA.
 DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004295; F:trypsin activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR003204; Cyt_C_ox5A.
 DR InterPro: IPR001545; Gly_hormoneB.
 DR InterPro: IPR001354; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR InterPro: IPR009003; Pept_Ser_Cys.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SMO0020; Tryp_SPC; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
 DR PROSITE: PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 KW Hydrolase; Protease; Serine protease.
 SO SEQUENCE 376 AA; 41434 MW; 64F403CAD505CD33 CRC64;

Query Match 89.5%; Score 34; DB 2; Length 376;
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLXDDLLEA 9
 Db 245 VLMDLLEA 253

RESULT 12

O27025 PRELIMINARY; PRT; 616 AA.
 ID O27025;
 AC O27025;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Type I restriction modification enzyme, subunit M.
 GN OrderedLocustNames=MT1942;
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxId=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Saefer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-T., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155 (1997).
 DR EMBL: AE000868; AAB85440.1; --
 DR PIR: C69226; C69226.

DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO: GO:0009007; F:site-specific DNA-methyltransferase (adenin. . .); IEA.
 DR GO: GO:0009306; P:DNA methylation; IEA.
 DR GO: GO:0006304; P:DNA modification; IEA.
 DR InterPro: IPR003665; Methylase_M.
 DR InterPro: IPR002296; N12N6_mtfase.
 DR InterPro: IPR003356; N6_DNA_Mcase.
 DR InterPro: IPR002052; N6_Mcase.
 DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam: PF02506; Methylase_M; 1.
 DR Pfam: PF02384; N6_Mcase; 1.
 DR PRINTS: PR00507; N12N6MTFASB.
 DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
 KW Complete proteome.
 SO SEQUENCE 616 AA; 71715 MW; 06D4F1076A5D5BAD CRC64;

Query Match 89.5%; Score 34; DB 2; Length 616;
 Best Local Similarity 77.8%; Pred. No. 2.7e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLXDDLLEA 9
 Db 465 VLMDLLEA 473

RESULT 13
 O94256 PRELIMINARY; PRT; 174 AA.
 ID O94256;
 AC O94256;
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Saposin-like protein family protein 19.
 GN Name=spp-19;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodietinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RS SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Wilson R.;

"Genome sequence of the nematode C. elegans: a platform for

```

RT Investigating biology. The C. elegans Sequencing Consortium.",
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Geisel C., Bradshaw H.,
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Waterston R.,
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Wilson R.,
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64849; AAC8055.1; -.
DR PIR; C89075; C89075.
DR WormPep; K04A8.9; CEI1716.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR011001; Saposin_Like.
DR SMART; SM00741; SapB; 1.
SQ SEQUENCE 174 AA; 19210 MW; A6303CF6383BBBD9 CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 174;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDLLEA 9
DB 129 VLXDLLEA 137

RESULT 14
ID Q82P65 PRELIMINARY; PRT; 185 AA.
AC Q82P65;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative TetR-family transcriptional regulator.
GN Ordered locus names=SAV1068;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinoe M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RL "Complete genome sequence and comparative analysis of the industrial
RL microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; AP005025; BAC68778.1; -.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001241; DNA_topoisom1.
DR InterPro; IPR009057; Homeodomain_Like.

```

```

DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS00177; TOPOISOMERASE_II; UNKNOWN 1.
KW Complete proteome; DNA-binding; Transcription regulation.
SQ SEQUENCE 185 AA; 19715 MW; 36B80407CDD54C30 CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 185;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDILLEA 9
DB 171 LXDILLEA 178

RESULT 15
ID Q81SY5 PRELIMINARY; PRT; 291 AA.
AC Q81SY5; O6KJB0;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Transcriptional regulator, Weir family.
GN Ordered locus names=BA1509; ORF names=GBA1509;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Bailly L.W., Paulsen I.T.,
RA Nelson K.E., Tetteijn H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapfel E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beaman M.J., Dodson R.O., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.F.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017028; AAP25446.1; -.
DR EMBL; AB017334; AAT35306.1; -.
DR TIGR; BA1509; -.
DR GO; GO:0005623; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011007; B12_binding.
DR InterPro; IPR000551; HTH_MeRR.
DR InterPro; IPR009061; Putativ_DNA_bind.
DR Pfam; PF00376; Weir; 1.
DR SMART; SM00422; HTH_MeRR; 1.
DR PROSITE; PS50937; HTH_MeRR_2; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 291 AA; 33164 MW; D96019E1884B142D CRC64;

Query Match
Best Local Similarity 86.8%; Score 33; DB 2; Length 291;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```
QY      1 VLXDDLEA 9
          |||||:|
Db      94 VLVDLLQA 102
```

Search completed: December 29, 2004, 23:12:51
Job time : 192 secs

THIS PAGE BLANK (USPTO)